Volume 46 *October 1960* *Number 10*

Proceedings of the
National Academy of Sciences
of the United States of America

The Proceedings of the National Academy of Sciences

OFFICERS OF THE ACADEMY

DETLEV W. BRONK
President

FARRINGTON DANIELS
Vice President

H. P. ROBERTSON
Foreign Secretary

HUGH L. DRYDEN
Home Secretary

L. V. BERKNER
Treasurer

EDITORIAL BOARD OF THE PROCEEDINGS

SAUNDERS MAC LANE
Chairman

HUGH L. DRYDEN, *Home Secretary*
L. V. BERKNER, *Treasurer*
H. P. ROBERTSON, *Foreign Secretary*
EDWIN B. WILSON, *Managing Editor*

GREGORY BREIT
HARRISON BROWN
S. CHANDRASEKHAR
ERNST CLOOS
THOMAS FRANCIS, JR.
C. H. GRAHAM
H. W. MAGOUN
HARLOW SHAPLEY
PAUL WEISS

Subscription price is $12.50 for 1 year, $24.00 for 2 years, and $34.50 for 3 years; the price of a single issue is $1.50. Make all remittances payable to the National Academy of Sciences in U.S. currency or its equivalent.

Editorial correspondence should be addressed to the PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 2101 Constitution Avenue, Washington 25, D. C.

Business correspondence should be addressed to the Printing and Publishing Office, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C.

Subscribers are requested to notify the Printing and Publishing Office of the Academy and their local postmaster immediately of change of address. Notices to the Academy should provide both the old and the new address.

Microfilms of complete volumes of this journal are available to regular subscribers only and may be obtained at the end of the volume year from University Microfilms, 313 N. First Street, Ann Arbor, Michigan.

Second-class postage paid at Easton, Pennsylvania. This publication is authorized to be mailed at the special rates of postage prescribed by Section 132.122.

PRINTED IN THE U. S. A.

THE PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES
is published monthly by
THE NATIONAL ACADEMY OF SCIENCES

Proceedings of the
NATIONAL ACDEMY OF SCIENCES

Volume 46 · Number 10 · October 15, 1960

*ON THE MECHANISM OF ACTION OF THE ANTIDIURETIC HORMONE (VASOPRESSIN)**

By Conrad T. O. Fong,† Lawrence Silver, David R. Christman, and Irving L. Schwartz

THE MEDICAL RESEARCH CENTER OF BROOKHAVEN NATIONAL LABORATORY, AND THE NEW YORK UNIVERSITY COLLEGE OF MEDICINE,

Communicated by Donald D. Van Slyke, August 5, 1960

Evidence available at present indicates that a primary function of the peptide hormones, vasopressin and insulin, is to influence membrane permeability. The chemical basis for this action is ill defined in spite of considerable knowledge concerning the structure and physiological action of these hormones. In our previous investigation findings on the interaction of vasopressin with kidney receptor sites indicated the formation of a covalent bond resulting from a thiol-disulfide exchange reaction.[1] The present paper substantiates this finding and suggests a plausible mechanism of action for the antidiuretic hormone.‡

Materials.—Preparation of tritiated vasopressin (H^3AVP): In our hands, tritiation by the usual Wilzbach method gave low yields and unpredictable specific activities. Hence we resorted to a modified Wilzbach procedure, employing an electric discharge for tritium labeling.[2]

AVP§ obtained from chromatographing beef pitressin powder[3] was introduced into a cup-shaped electrode attached to one-half of a cylindrical pyrex cell. After sealing, the cell was evacuated and tritium gas was allowed to flow in (pressure less than 5 mm Hg). An electric discharge was produced by connecting the electrodes to an AC transformer in series with an autotransformer. The cell was re-evacuated to remove excess tritium and other gases and the tritiated H^3AVP was dissolved in dilute acetic acid and lyophilized. This latter process was repeated three times and the lyophilized protein was then purified twice by chromatography through a CM cellulose column according to the method of Ward and Guilleman.[3] H^3AVP with a specific activity of approximately 400 μc/mg and a biological activity of about 400 pressor units/mg[4] was obtained, indicating a high degree of purification (Fig. 1).

Preparation of acetyl LVP: Acetyl LVP was prepared by reacting LVP (prepared from pitressin powder according to the method of Ward and Guilleman[3]) with acetic anhydride in glacial acetic acid. The acetyl derivative was subsequently purified by chromatographic separation using a CM cellulose column.

Iodination of AVP with tracer I^{131}: Iodination of the hormone was carried out by the usual KI_3 method using 0.2 *M* ammonium hydroxide in place of bicarbonate.[5] The IAVP was purified through a CM cellulose column.

The Proceedings of the National Academy of Sciences

OFFICERS OF THE ACADEMY

DETLEV W. BRONK
President

FARRINGTON DANIELS
Vice President

H. P. ROBERTSON
Foreign Secretary

HUGH L. DRYDEN
Home Secretary

L. V. BERKNER
Treasurer

EDITORIAL BOARD OF THE PROCEEDINGS

SAUNDERS MAC LANE
Chairman

HUGH L. DRYDEN, *Home Secretary*
L. V. BERKNER, *Treasurer*
H. P. ROBERTSON, *Foreign Secretary*
EDWIN B. WILSON, *Managing Editor*

GREGORY BREIT
HARRISON BROWN
S. CHANDRASEKHAR
ERNST CLOOS
THOMAS FRANCIS, JR.
C. H. GRAHAM
H. W. MAGOUN
HARLOW SHAPLEY
PAUL WEISS

Subscription price is $12.50 for 1 year, $24.00 for 2 years, and $34.50 for 3 years; the price of a single issue is $1.50. Make all remittances payable to the National Academy of Sciences in U.S. currency or its equivalent.

Editorial correspondence should be addressed to the PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 2101 Constitution Avenue, Washington 25, D. C.

Business correspondence should be addressed to the Printing and Publishing Office, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C.

Subscribers are requested to notify the Printing and Publishing Office of the Academy and their local postmaster immediately of change of address. Notices to the Academy should provide both the old and the new address.

Microfilms of complete volumes of this journal are available to regular subscribers only and may be obtained at the end of the volume year from University Microfilms, 313 N. First Street, Ann Arbor, Michigan.

Second-class postage paid at Easton, Pennsylvania. This publication is authorized to be mailed at the special rates of postage prescribed by Section 132.122.

PRINTED IN THE U. S. A.

THE PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES
is published monthly by
THE NATIONAL ACADEMY OF SCIENCES

Proceedings of the
NATIONAL ACDEMY OF SCIENCES

Volume 46 · Number 10 · October 15, 1960

*ON THE MECHANISM OF ACTION OF THE ANTIDIURETIC HORMONE (VASOPRESSIN)**

By Conrad T. O. Fong,† Lawrence Silver, David R. Christman, and Irving L. Schwartz

THE MEDICAL RESEARCH CENTER OF BROOKHAVEN NATIONAL LABORATORY, AND THE NEW YORK UNIVERSITY COLLEGE OF MEDICINE,

Communicated by Donald D. Van Slyke, August 5, 1960

Evidence available at present indicates that a primary function of the peptide hormones, vasopressin and insulin, is to influence membrane permeability. The chemical basis for this action is ill defined in spite of considerable knowledge concerning the structure and physiological action of these hormones. In our previous investigation findings on the interaction of vasopressin with kidney receptor sites indicated the formation of a covalent bond resulting from a thiol-disulfide exchange reaction.[1] The present paper substantiates this finding and suggests a plausible mechanism of action for the antidiuretic hormone.‡

Materials.—Preparation of tritiated vasopressin (H^3AVP): In our hands, tritiation by the usual Wilzbach method gave low yields and unpredictable specific activities. Hence we resorted to a modified Wilzbach procedure, employing an electric discharge for tritium labeling.[2]

AVP§ obtained from chromatographing beef pitressin powder[3] was introduced into a cup-shaped electrode attached to one-half of a cylindrical pyrex cell. After sealing, the cell was evacuated and tritium gas was allowed to flow in (pressure less than 5 mm Hg). An electric discharge was produced by connecting the electrodes to an AC transformer in series with an autotransformer. The cell was re-evacuated to remove excess tritium and other gases and the tritiated H^3AVP was dissolved in dilute acetic acid and lyophilized. This latter process was repeated three times and the lyophilized protein was then purified twice by chromatography through a CM cellulose column according to the method of Ward and Guilleman.[3] H^3AVP with a specific activity of approximately 400 μc/mg and a biological activity of about 400 pressor units/mg[4] was obtained, indicating a high degree of purification (Fig. 1).

Preparation of acetyl LVP: Acetyl LVP was prepared by reacting LVP (prepared from pitressin powder according to the method of Ward and Guilleman[3]) with acetic anhydride in glacial acetic acid. The acetyl derivative was subsequently purified by chromatographic separation using a CM cellulose column.

Iodination of AVP with tracer I^{131}: Iodination of the hormone was carried out by the usual KI_3 method using 0.2 *M* ammonium hydroxide in place of bicarbonate.[5] The IAVP was purified through a CM cellulose column.

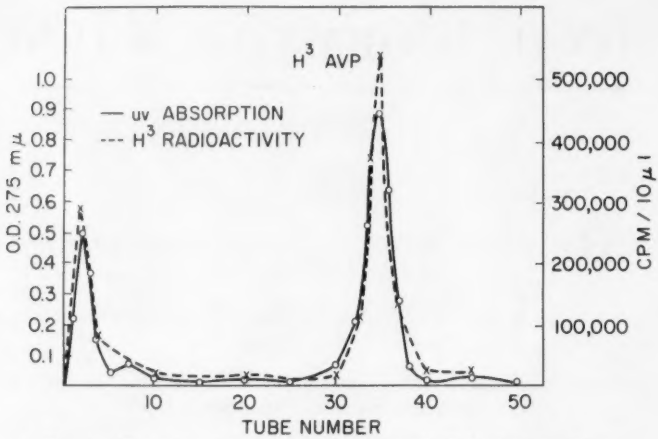


FIG. 1.—Chromatography of H³AVP through CM cellulose column. —— u.v. curve; - - - - radioactivity curve.

Methods.—In vivo binding and separation of hormone-receptor proteins: The following experiment demonstrates in vivo combination of H³AVP and a kidney protein fraction. Partial dissociation of this complex was effected by treatment with cysteine, mercaptoethylamine, and thioglycollate.

Seven rats were hydrated by stomach tube and anesthetized with dilute ethanol by mouth. Approximately 5 μg of H³AVP in 0.5 ml saline were injected into each rat over a 5-min period via a cannulated external jugular vein. Ten minutes after the onset of the injection of H³AVP, at the peak of antidiuretic activity, the kidneys were quickly excised and dropped into liquid nitrogen. These were homogenized with nine volumes of 30 per cent ice-cold sucrose,[6] saturated with NEM to bind sulfhydryl groups and thereby block any interchange reaction, and centrifuged for 20 min at 400 $\times$ g and 5°C. The residue was suspended in several volumes of 30 per cent sucrose and recentrifuged at 400 $\times$ g. The precipitate was then rehomogenized in saline containing NEM and recentrifuged at 4000 $\times$ g for one half-hr. The pellet, containing cell membrane debris and nuclei, was suspended in water, dialyzed against cold tap water and electrodialyzed to remove ionic bonded H³AVP. After electrodialysis acetone was added to the aqueous suspension to a concentration of about 90 per cent and the precipitated proteins were centrifuged and dried in vacuo.

Equal portions of these fractionated kidney proteins were shaken overnight at 37°C with the following reagents made up in 0.1 M sodium bicarbonate, adjusted to pH 8.0: (a) bicarbonate control; (b) 0.1 M cysteine; (c) 0.1 M β-mercaptoethylamine; (d) 0.1 M thioglycollate. The treated tissues were centrifuged, in sequence, with water and acetone and dried in vacuo. The dried kidney proteins thus obtained were prepared for tritium analysis by a modified Wilzbach combustion method,[7, 8] using gas proportional counting.

Results.—In a series of fractionations higher specific activity of H³AVP-kidney protein complex was obtained by partial sucrose fractionation than by the original procedure.[1] Table 1 shows a typical experiment in which sucrose fractionation

TABLE 1

H³AVP CONTENT OF SUCROSE FRACTIONATED KIDNEY PROTEINS TREATED WITH VARIOUS THIOL COMPOUNDS IN BICARBONATE AT pH 8

Treatment	Disintegrations per min per mg protein		Average dpm per mg	% H³AVP released
0.1 *M* sodium bicarbonate (control)	42.91	45.09	43.50	...
0.1 *M* cysteine	22.49	22.39	22.45	48.5
0.1 *M* β-mercaptoethylamine	19.81	19.34	19.57	54.5
0.1 *M* thioglycolate	22.61	22.91	22.76	47.7

TABLE 2

H³AVP CONTENT OF FRACTIONATED KIDNEY PROTEINS (1500 × *g*) AFTER TREATMENT WITH THIOL COMPOUNDS IN 8 *M* UREA pH 8

Treatment	Disintegrations per min per mg protein			Average dpm per mg	% H³AVP released
Control	4.8	3.3	4.7	4.3	...
2% Cysteine		2.3		2.3	46
2% β-Mercaptoethylamine		2.3		2.3	46
2% Thioglycolate	2.23		1.47	1.8	58

was employed and Table 2 shows the results from a fractionation procedure essentially as previously described using 8 *M* urea and sodium hydroxide instead of bicarbonate in the thiol-compound treatment. Examination of these tables shows a quantitative difference of bonded H³AVP following thiol-compound treatment of the protein. The release of approximately 46-58% H³AVP indicates a cleavage of a sulfur-sulfur bond by mercaptide ions.

Discussion.—Our experiments suggest a sulfur-sulfur bond in the hormone-receptor interaction. The basis for proposing a specific thiol-disulfide exchange reaction and the evidence in support thereof are:

1. The investigations of Ames and Van Dyke[9] on pitressin and our observations on vasopressin (unpublished results) showed that the loss of antidiuretic activity paralleled the reduction of the hormone disulfide.

2. The studies of Cafruny, Carhart, and Farah[10] on the effects of various hormones on the thiol concentration in kidney cells from hypophysectomized rats showed that treatment with pitressin produced a diminution of sulfhydryl groups in the convoluted portions of the proximal and distal tubules. No changes in thiol concentration in the same kidney tubule sections were observed with growth hormone (a disulfide hormone) and other hormones (thyroxin and thyrotropin). This finding suggests a specific hormone-receptor interaction in the distal convoluted tubule, a segment in which vasopressin-induced water reabsorption is known to occur. The similar finding in the proximal convoluted tubule raises the possibility of a similar hormone-receptor interaction at this site in the nephron, although the functional significance of a proximal permeability change is not apparent.

3. The release of reduced tritiated hormone from hormone-receptor kidney proteins after treatment with thiol compounds (Table 1) indicates a cleavage of a disulfide bond.

4. Toad bladders (*Bufo marinus*) treated with thiol-blocking agents, such as NEM, *p*-chloromercuribenzoate and methylmercuricbromide, showed considerably less disulfide-linked H³AVP than did untreated bladders.[11]

The interaction of the antidiuretic hormone with its receptor may require a specificity dictated by stereochemical and electrostatic effects.

With the structure of arginine vasopressin, CyS·Tyr·Phe·Glu(NH_2)·Asp(NH_2)·Cys·Pro·Arg·Gly(NH_2),[12, 13] and all the above considerations in mind, a plausible mechanism for the action of the antidiuretic hormone is:

(*a*) The hormone is attracted to its receptor site by electrostatic interactions of opposite charges, i.e., the positively charged groups of the hormone and the negatively charged groups of the receptor. At short ranges hydrogen bonding, through the tyrosyl hydroxyl group and amide groups of glutamine, asparagine, and glycinamide, emerges as a prominent attractive force. The proper interactions between attractive and repulsive forces bring about an alignment of the hormone with its receptor site.

(*b*) The thiol-disulfide exchange reaction, which is a nucleophilic attack of the receptor mercaptide ion on the hormone disulfide, takes place. This results in a hormone-receptor disulfide. Subsequently, a series of sulfhydryl-disulfide reactions induces alterations of the tertiary structure of membrane proteins. This chain reaction brings about conformational changes in the protein components of the diffusion barrier. The resultant molecular sieve-like structure allows for an increase in the flux of water and certain solutes through the partially "loosened" membrane.

(*c*) The hormone-receptor disulfide bond is ultimately cleaved by a reductase in conjunction with one or more other enzymes.[14, 15] Such a reverse reaction in which the receptor sulfhydryl group is regenerated restores the membrane to its original state and completes the cycle (Fig. 2).

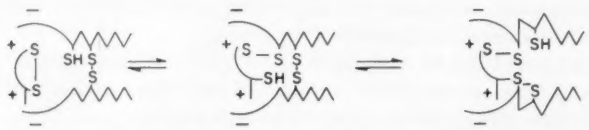


FIG. 2.—Schematic mechanism of antidiuretic hormone action.

A few preliminary experiments were designed to test several points of this hypothesis. These included the elimination of the positively charged groups by acetylation of LVP. This derivative gave an antidiuretic activity (in the rat) of approximately $^1/_{400}$ of the original. Iodination of AVP and LVP resulted in derivatives manifesting about $^1/_{100}$ the antidiuretic activities of the original hormones. This suggests a reduction in hydrogen bonding through the tyrosyl hydroxyl group. These data are consistent with the postulated hypothesis. It is interesting to note the presence of a terminal prolyl peptide chain in a number of hormones (oxytocin, insulin, melanocyte-stimulating hormone): such a chain may contribute to molecular specificity.[16, 17] An examination of structural analogs of vasopressin[18–21] and the correlation of their antidiuretic activities in the dog and rat[22] is consistent with this hypothesis and implies the necessity for a correct molecular fit of the hormone to its receptor. In this mechanism the disulfide of the hormone is considered the essential functional group. A review is available on the role of sulfhydryl-disulfide interchange reactions in biological processes concerned with the phenomena of protein denaturation, blood clotting, and mitosis.[23]

It is interesting that one of the prime functions of insulin is to alter the permeability of insulin-sensitive membranes to sugars.[24] Investigations of Lindley[25] and Fraenkel-Conrat[26] have indicated that the intrachain disulfide bond of insulin is preferentially cleaved on partial reduction and hormonal activity is thus lost. This suggests that a thiol-disulfide exchange reaction similar to that postulated for vasopressin may be involved.

Summary.—Evidence is presented to indicate that a covalent bond is formed in the interaction of vasopressin and its receptor. The disulfide bond thus formed is a result of a thiol-disulfide exchange reaction. A mechanism of action for the antidiuretic hormone is suggested which assumes that this exchange reaction underlies the increased passive transport of water induced by vasopressin.

We gratefully acknowledge the interest and support of Dr. L. Thomas, Dr. A. B. Gutman and the J. M. Kaplan Foundation, and the assistance of Miss M. A. Schoessler.

* This research was supported in part by a grant from the National Science Foundation and in part by the U. S. Atomic Energy Commission.

† Inquiries should be sent to the author at Brookhaven National Laboratory, Upton, New York.

‡ AVP and LVP are referred to by synonyms, "vasopressin" and "antidiuretic hormone."

§ Abbreviations used are: AVP, H^3AVP, IAVP, arginine vasopressin, tritiated arginine vasopressin, and iodinated vasopressin, respectively; LVP, lysine vasopressin; CM cellulose, carboxymethyl-cellulose; NEM, N-ethylmaleimide.

[1] Fong, C. T. O., I. L. Schwartz, E. A. Popenoe, L. Silver, and M. A. Schoessler, *J. Am. Chem. Soc.*, **81**, 2592 (1959).

[2] Dorfman, L. M., and K. E. Wilzbach, *J. Phys. Chem.*, **63**, 799 (1959).

[3] Ward, D. N., and R. Guilleman, *Proc. Soc. Exp. Biol. Med.*, **96**, 568 (1957).

[4] Thorn, N. A., *J. Exp. Med.*, **105**, 585 (1957).

[5] Hughes, W. L., and R. Straessle, *J. Am. Chem. Soc.*, **72**, 452 (1950).

[6] Kuff, E. L., G. H. Hogeboom, and A. J. Dalton, *J. Biophys. Biochem. Cytol.*, **2**, 33 (1956).

[7] Wilzbach, K. E., L. Kaplan, and W. G. Brown, *Science*, **118**, 522 (1953).

[8] Christman, D. R., *Chemist Analyst*, **46**, 5 (1957).

[9] Ames, R. G., and H. B. Van Dyke, *Proc. Soc. Exp. Biol. Med.*, **76**, 576 (1951).

[10] Cafruny, E. J., E. Carhart, and A. Farah, *Endocrinol.*, **61**, 143 (1957).

[11] Schwartz, I. L., H. Rasmussen, M. A. Schoessler, C. T. O. Fong, and L. Silver, *J. Clin. Invest.*, **39**, 1026 (1960).

[12] Popenoe, E. A., and V. du Vigneaud, *J. Biol. Chem.*, **206**, 353 (1954).

[13] du Vigneaud, V., H. C. Lawler, and E. A. Popenoe, *J. Am. Chem. Soc.*, **75**, 4880 (1953).

[14] Dicker, S. E., and A. L. Greenbaum, *J. Physiol.*, **141**, 107 (1958).

[15] Dicker, S. E., in *Polypeptides Which Affect Smooth Muscles and Blood Vessels*, ed. M. Schachter (London: Pergamon Press, 1960), p. 79.

[16] Ressler, C., *Proc. Soc. Exp. Biol. Med.*, **92**, 725 (1956).

[17] Pauling, L., and R. B. Corey, *Fortschr. Org. Chem. Naturst*, **XI**, 180 (1954).

[18] Katsoyannis, P. G., and V. du Vigneaud, *Arch. Biochem. & Biophys.*, **78**, 555 (1958).

[19] Katsoyannis, P. G., *J. Am. Chem. Soc.*, **79**, 109 (1957).

[20] Katsoyannis, P. G., and V. du Vigneaud, *J. Biol. Chem.*, **233**, 1352 (1958).

[21] Boissonnas, R. A., in *Polypeptides Which Affect Smooth Muscles and Blood Vessels*, ed. M. Schachter (London: Pergamon Press, 1960), p. 7.

[22] Munsick, R. A., W. H. Sawyer, and H. B. van Dyke, *Endocrinol.*, **66**, 860 (1960).

[23] Jensen, E. V., *Science*, **130**, 1319 (1959).

[24] Levine, R., and M. S. Goldstein, *Rec. Progr. Horm. Res.*, **11**, 343 (1955).

[25] Lindley, H., *J. Am. Chem. Soc.*, **77**, 4927 (1955).

[26] Fraenkel-Conrat, J., and H. Fraenkel-Conrat, *Biochim. et Biophys. Acta*, **5**, 89 (1950).

*STUDIES ON THE MECHANISM OF ACTION OF VASOPRESSIN**

By Howard Rasmussen†, Irving L. Schwartz, Mary Anne Schoessler, and Gerty Hochster

THE ROCKEFELLER INSTITUTE AND THE MEDICAL RESEARCH CENTER OF BROOKHAVEN NATIONAL LABORATORY

Communicated by Lyman C. Craig, August 4, 1960

In spite of the work of du Vigneaud and his collaborators on the isolation, characterization, and synthesis of the neurohypophyseal hormones, vasopressin, and oxytocin,[1] relatively little is known concerning their basic mechanism of action. Studies by Ussing, Leaf, Bentley, and Sawyer indicate that these agents change the permeability of amphibian membranes to water, urea, and sodium ions.[2–11] One of the most simple structurally of these membranes is the urinary bladder of the toad, *Bufo Marinus*, which consists of a single layer of epithelial cells supported by a minimal amount of connective tissue, smooth muscle, and blood vessels.[5] Work done with this membrane indicates that the bladder wall is relatively impermeable to water and urea under resting conditions, and that the neurohypophyseal hormones increase dramatically the permeability.[6–8] A simple *in vitro* system has been devised to measure this action of the hormones.[7] Such a system has been utilized in the present study in an effort to define the mechanism of action of the neurohypophyseal hormone, arginine vasopressin. Two aspects of the problem have been investigated. The results indicate that the change in permeability induced by arginine vasopressin (AVP) is not directly dependent upon oxidative metabolism and that the attachment of the hormone to the bladder involves, in all likelihood, the formation of a mixed disulfide.

Effects of Oxygen, Substrate, and Temperature upon the Action of Vasopressin.—The first question to be studied was whether or not the hormones catalyzed a reaction which is linked directly to the energy yielding reactions of the cell.

In the studies to be reported, the method of Bentley[7, 8] was employed. Either an amphibian Ringer (Na^+ – 112 m*M*, K^+ – 3.2 m*M*, Cl^- – 114 m*M*, HCO_3^- – 2.0 m*M*, $HPO_4^=$ – 1.2 m*M*, Mg^{++} – 1.0 m*M*, $SO_4^=$ – 1.0 m*M*, Ca^{++} – 2.0 m*M*, mOs*M* – 220, pH – 7.36) or a similar solution with the sodium ion replaced by choline ion was used. Hypotonic Ringer solutions were prepared by dilution 1 to 5 with distilled water. (54 mOsM). The osmolar concentrations of the solutions were measured in the Fiske Osmometer. The toads, weighing approximately 200 gm, were kept well-hydrated but not fed. They were pithed, and the bladders dissected out immediately before use. The urinary bladder is a bilobed structure, which was removed in two parts, one half usually serving as control for the second. Each lobe was mounted upon the tip of a hollow glass rod which extended through a rubber stopper. A wire loop was attached to the rod. After the bladder was mounted (mucosal side in) 2 ml of hypotonic Ringer were added through the hollow glass tube, and the mounted bladder suspended in 20 ml of isotonic Ringer (serosal solution) through which was bubbled continuously either air, nitrogen, or oxygen. The mounted bladder assembly was removed from the serosal solution after a predetermined time and weighed without blotting to ±2 mg. The weight loss was recorded as mg/min. Under control conditions with $^1/_5$ Ringer's solution inside

(mucosal) and Ringer's solution outside (serosal) the weight loss averaged approximately 0.5–1 mg/min, whereas following the addition of arginine vasopressin (0.0125γ/ml) to the serosal solution the weight loss increased to 10–20 mg/min. The mucosal and serosal solutions were changed at the end of each 20 min period. The data reported are the mean of at least six separate experiments. The results were usually quite consistent ($\pm 15\%$).

In Figure 1 is illustrated a representative experiment which shows the effect of 0.0125γ/ml of arginine vasopressin (AVP) upon the permeability of the bladder in sodium and choline solutions. The rate of weight loss of bladders kept in choline is approximately 50 per cent of that in sodium, although during the control period the bladder placed in the choline Ringer was more permeable than the one in sodium

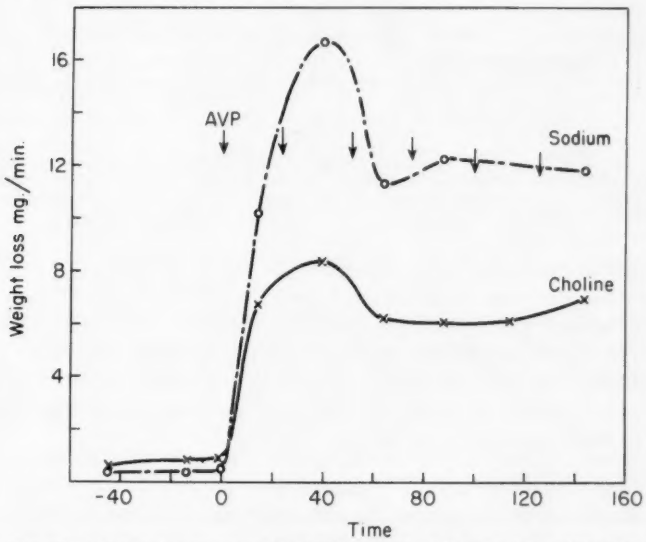


FIG. 1.—The effect of arginine vasopressin (AVP), 0.0125γ/ml, upon the weight loss of the isolated urinary bladder of the toad in vitro in sodium and in choline Ringer. The arrows indicate the points at which both the mucosal and serosal solutions were changed. When the serosal solution was changed, a new dose of AVP was added, so that the hormone concentration was maintained relatively constant. Time is in min.

Ringer. The striking increase in water movement is a passive process because when the osmotic gradient was reversed by the addition of sufficient mannitol to render the mucosal solution hypertonic to the serosal bath, the direction of the water movement was also reversed. A similar observation has been reported by Sawyer who worked with the bullfrog bladder.[11]

In Figure 2 are summarized the results of a series of experiments which show that the rate of change of permeability was independent of the oxygen and glucose concentrations. Choline Ringer was used to eliminate variations in sodium transport and the attendant changes in oxygen consumption. The results shown in Figure 2 were obtained from experiments carried out at 8.5° but similar results were noted at 17°.

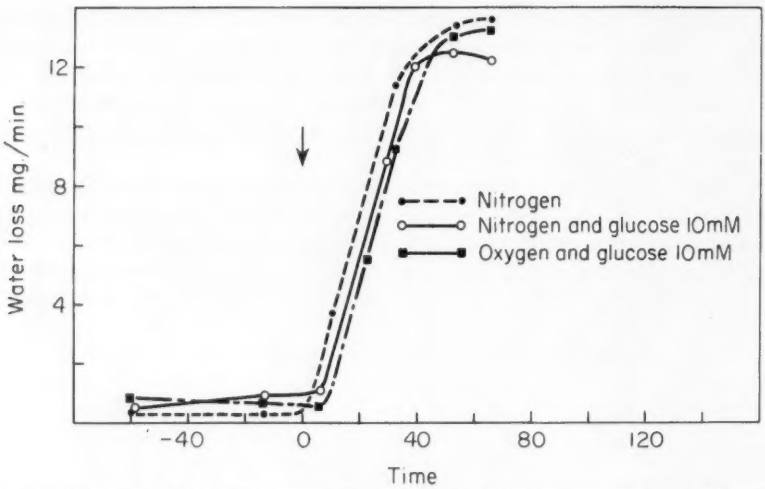


FIG. 2.—The effect of changes in oxygen and glucose concentration upon the rate of change of bladder permeability (as measured by weight loss) induced by AVP (0.0125γ/ml) in choline Ringer. The experiments were carried out at 8.5°. Either oxygen or nitrogen was bubbled through the serosal choline solution in which substrate (glucose) was either present or absent.

In Table 1 are recorded data concerning the effects of arginine vasopressin on the oxygen consumption of the isolated toad bladder in choline Ringer. These studies, which were carried out using classical manometric techniques, indicate that there is no change in oxygen consumption upon the addition of hormone in choline Ringer solution. By contrast, Leaf *et al.* have shown that the addition of hormone to the bladder in sodium Ringer results in an increased oxygen consumption,[17] presumably because of increased sodium transport.

TABLE 1

THE EFFECT OF VARIOUS AGENTS UPON THE OXYGEN CONSUMPTION OF THE ISOLATED TOAD BLADDER AT 25°

Agent	Oxygen Consumption, μl O_2/mg dry wt/hr	
	Sodium Ringer	Choline Ringer
Control	1.07 ± 0.11	0.68 ± 0.16
AVP	1.40 ± 0.17	0.65 ± 0.15
10^{-4} NEM serosal	0.29 ± 0.07	0.22 ± 0.07
10^{-4} NEM mucosal, first 30 min	1.18 ± 0.09	....
10^{-4} NEM mucosal, 30–60 min	0.76 ± 0.06	....
pH 6.6 serosal	1.08 ± 0.14	....

The effects of change of temperature upon the action of the hormone was also studied. Arginine vasopressin increased the water permeability of the bladder kept at 2.5°, under nitrogen, in choline Ringer, and in the presence of 10^{-2} KCN. The most striking effect of a change in temperature was upon the rate of initiation of the change in permeability. For example, at 17° there appeared to be little or no lag in the opening of the permeability barrier whereas at 8.5°, a delay of 10–12 min, and at 2.0° a delay of 20–25 min were noted. From the difference in slopes of the curves obtained at 8.5° and at 17°, a Q^{10} for the rate of change of permeability was calculated to be 1.66.

Bentley[7] has shown that the permeability of the bladder to water returns to control values when the hormone is washed out of the solution bathing the serosal surface or if that solution is made more acid than pH 6.8. As shown in Table 1 the oxygen consumption of bladders incubated at pH 6.6 is not significantly different from those incubated in the standard pH 7.6 solution.

These results indicate that in all likelihood the chemical reactions which control the reversible opening and closing of the permeability barrier of these cells, and which are regulated by the neurohypophyseal hormones, are not directly linked to the oxidative metabolism of these cells.

Actions of Sulfhydryl Agents.—Since the hormone contains a disulfide bond, which is essential for its action,[13] it is possible, as has already been suggested,[14, 15] that the primary interaction between hormone and receptor site involves a disulfide-sulfhydryl interchange reaction between sulfhydryl (SH) groups at the receptor site and the disulfide S-S linkage in the hormone molecule. To test this possibility, a number of the more common sulfhydryl reagents and sulfhydryl compounds were tested for their ability to augment or inhibit the action of vasopressin. These studies were all carried out at 17° with either choline or sodium Ringer. The results are summarized in Table 2. A number of general metabolic inhibitors had little effect upon AVP action when added to bladders immersed in choline solutions whereas a number of sulfhydryl reagents had a marked effect. For example, N-ethylmaleimide (NEM) added to the serosal solution in concentrations greater than 1×10^{-3} *M* inhibited completely the effect of vasopressin. At a concentration of 10^{-4} *M*, inhibition was only partial.

The hormones oxytocin and vasotocin were found to increase the permeability to water in the toad bladder, in confirmation of Sawyer's observations upon the frog bladder.[11] Vasotocin was several hundred times more potent than vasopressin. Pretreatment of bladders with NEM or *p*-hydroxymercuribenzoate inhibited the action of both vasotocin and oxytocin as well as of vasopressin.

NEM was also able to fix the bladder in a more permeable state after vasopressin had altered the permeability. As illustrated in Figure 3, addition of arginine vasopressin to the serosal solution caused an increase in permeability which was reversed when the AVP was removed. If, however, 10^{-3} *M* NEM was added before the AVP was washed out, the bladder remained open after the AVP had been removed. Furthermore, as shown in Figure 4, this effect of NEM was not completely reversed by bringing the pH of the serosal solution to 6.3, whereas the effect of AVP alone was completely reversed by this procedure.

As may be seen from Table 2, a number of organic mercurials produced the same effect as NEM, but neither iodacetamide, nor a number of sulfhydryl or disulfide containing compounds altered significantly the action of AVP.

It has been proposed that the permeability barrier, and, therefore, the site of action of AVP, is at or near the mucosal surface of the epithelial cells of the bladder.[16] Yet it has been shown that AVP is not active when added to the mucosal surface.[5] The current explanation of this fact is that the mucosal barrier is impermeable to a molecule the size of AVP, so that the latter is unable to penetrate the barrier and get to the site of action. Because NEM is a much smaller molecule, it seemed possible that it might act when added directly to the mucosal solution. Such a possibility was confirmed, for when either 10^{-3} *M* NEM or methylmercuric bromide

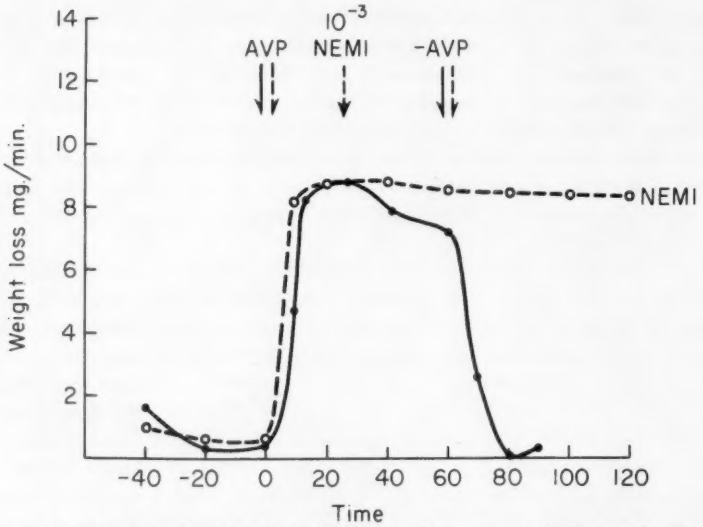


FIG. 3.—The effect of N-ethylmaleimide (NEM) upon the reversal of the increased permeability of isolated toad bladder induced by AVP in choline Ringer. The AVP was added to both sets of bladders at time zero, NEM was added to one set (----) after 25 min, then at 60 min the AVP was washed out of both sets. The permeability of the control bladders rapidly returned to normal, whereas those treated with NEM remained highly permeable.

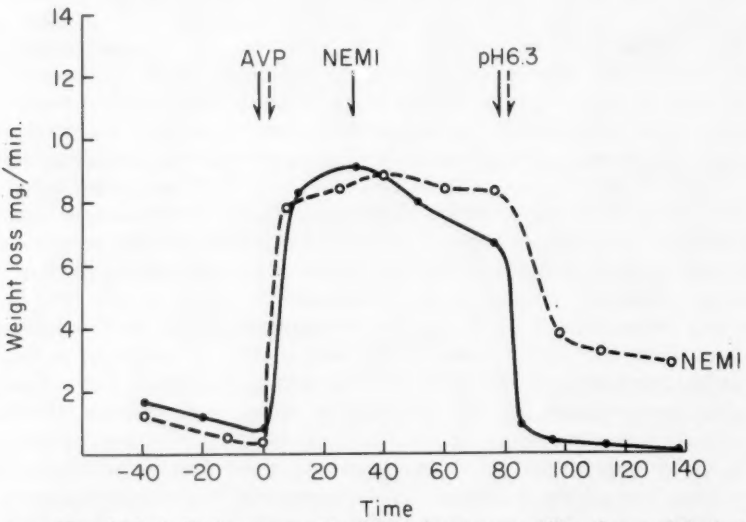


FIG. 4.—The effect of pH change upon the AVP-induced permeability change of the isolated toad bladder in choline Ringer. When the serosal pH was decreased from 7.3 to 6.3 the permeability of the AVP treated bladders returned to control values, whereas bladders treated with AVP and then NEM were maintained in a state of increased permeability even when the pH was lowered to 6.3.

TABLE 2

THE EFFECT OF VARIOUS COMPOUNDS UPON THE PERMEABILITY OF THE TOAD BLADDER AND UPON THE ACTION OF ARGININE VASOPRESSIN

Compound	Concentration	Permeability	AVP action
Serosal Surface			
N-ethylmaleimide	10^{-3}	...	Inhibition
N-ethylmaleimide	10^{-4}	...	Inhibition—(partial)
Methyl mercuric bromide	10^{-3}	... or I	Inhibition
p-hydroxymercuribenzoate	10^{-3}	...	Inhibition
p-chloromercuriphenyl sulfonic acid	10^{-3}	...	Inhibition
Mercaptoethylamine	10^{-3}	...	None
Reduced glutathione	10^{-3}	...	None
Oxidized glutathione	10^{-3}	...	None
Cystine	10^{-2}	...	None
Cysteine	10^{-2}	...	None
Iodoacetamide	10^{-2}	...	None
Sodium sulfite	10^{-2}	...	None
Potassium cyanide	10^{-3}	...	None*
Sodium fluoracetate	10^{-2}	...	None*
Dinitrophenol	10^{-3}	...	None*
Mucosal Surface			
N-ethylmaleimide	10^{-3}	... or SI	Inhibition
N-ethylmaleimide	10^{-3} for 2 min	I	Inhibition
N-ethylmaleimide	10^{-4}	I	Inhibition
N-phenylmaleimide	10^{-3}	...	Inhibition
N-phenylmaleimide	10^{-4}	I	Inhibition
Methyl merduric bromide	10^{-3}	... or SI	Inhibition
Methyl mercuric bromide	10^{-4}	I	Inhibition
Methyl mercuric bromide	10^{-6}	...	None
Iodoacetamide	10^{-2}	...	Inhibition—(partial)
Pencillamine	10^{-3}	...	None
Succinimide	10^{-3}	...	None
N-Bromosuccinimide	10^{-3}	...	Inhibition
Disodium versenate	10^{-2}	...	None
Potassium cyanide	10^{-3}	...	None
Sodium sulfite	10^{-3}	...	None
Mercaptoethanol	10^{-2}	...	None

Under the permeability column (...) indicates no change; (I) indicates a significant increase in permeability; and (SI) a slight increase.

* These agents have no effect on the action of AVP upon bladders immersed in choline Ringer, but do produce a slight inhibition of AVP action in sodium Ringer.

were added to the mucosal solution they completely inhibited the action of AVP added subsequently to the serosal solution. Iodoacetamide at the same concentration produced a partial inhibition of AVP action. When NEM, N-phenylmaleimide, or methylmercuric bromide (MMBr) were added to the mucosal surface in concentrations of 10^{-4} or 10^{-5} M, they induced a change in bladder permeability qualitatively similar to that seen after the addition of AVP to the serosal surface. Furthermore, the effect was more pronounced in sodium than in choline solutions (Fig. 5). The action of NEM 10^{-4} M was irreversible whereas that of methyl mercuric bromide was partially reversible.

If the mucosal surfaces of the bladders were bathed in 10^{-3} M NEM for 1.5–5 min and then washed free of this agent, their permeability increased subsequently as shown in Figure 6. The pattern observed was quite similar to that seen after the addition of 1.5×10^{-6} M AVP to the serosal surface (Fig. 1).

When 10^{-4} M NEM was added to the mucosal solution, it had little effect upon oxygen consumption during the first 30 min of incubation (Table 1) and depressed oxygen consumption only 25–30 per cent in the next 30-min period, whereas the

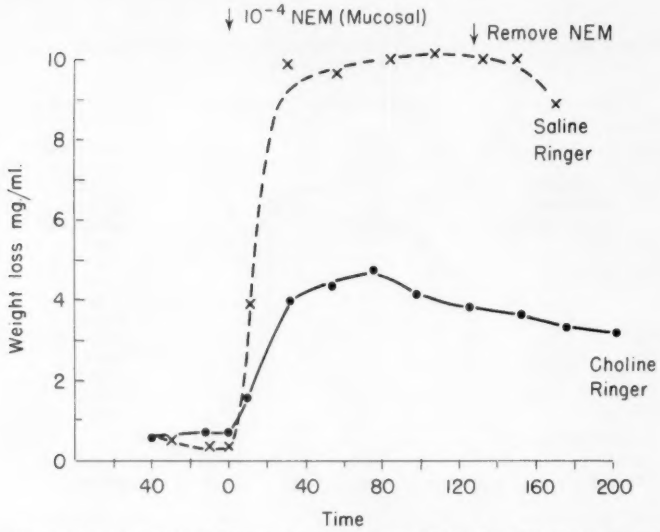


FIG. 5.—The effect of 10^{-4} NEM upon the weight loss of the isolated bladder in saline and in choline Ringer. The NEM was added to the mucosal at time zero. Compare with Figure 1.

same concentration of NEM added to the serosal solution depressed oxygen consumption 80–90 per cent within 15–20 min (Table 1). As illustrated in Figure 6, the maximal change in permeability induced by the addition of NEM to the mucosal

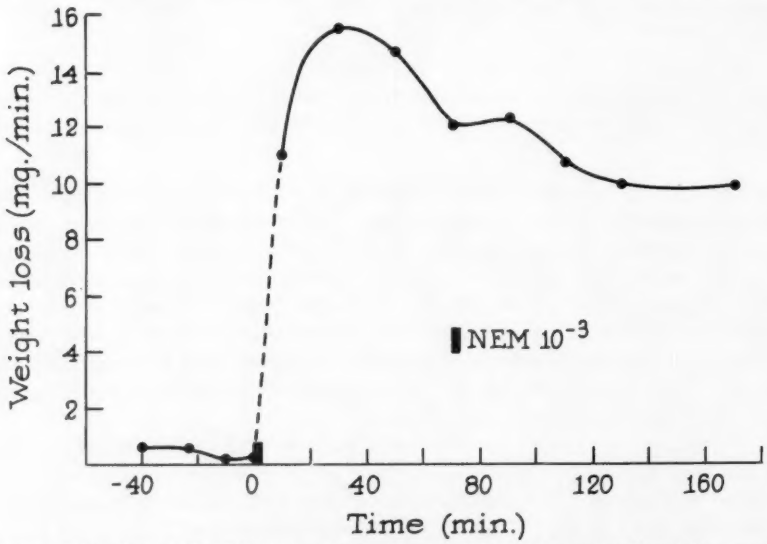


FIG. 6.—The effect of 10^{-3} *M* NEM added to the mucosal solution for 2 min and then washed out. Following this treatment the weight loss was followed in the usual manner with $^1/_5$ sodium Ringer solution inside (mucosal) and sodium Ringer outside. No AVP was added. Because of the washing out procedure there is a gap in the recording of the weight loss.

surface occurred within 20 min, a time at which there was no change in oxygen consumption. This suggests, as do the results in Figure 6, that the NEM was acting at or near the mucosal surface of the cell and not penetrating into the cell to any great extent. These results support Leaf's proposal[16] that the permeability barrier of these cells is at or near their mucosal surface.

In Table 2 are summarized the data concerning the effects of various other agents upon bladder permeability and AVP action when added to the mucosal surface. All of the compounds which were effective in either inhibiting AVP action or mimicing the effects of this hormone are well-known sulfhydryl reagents. N-bromosuccinimide, which inhibited the effect of AVP, is not widely used as a sulfhydryl reagent but undoubtedly reacts readily with sulfhydryl groups. It is noteworthy that succinimide, the saturated analog of maleimide, was completely ineffective in either changing bladder permeability or inhibiting the effect of AVP.

Discussion.—The data presented indicate that the primary reaction of arginine vasopressin with its receptor site in the amphibian bladder is not directly dependent upon metabolic energy. In addition, the data are consistent with the supposition that one of the primary interactions between the hormone and receptor involves the disulfide bond of the hormone and a sulfhydryl or related grouping in the receptor. It is proposed that under control conditions the permeability of the bladder is maintained by a protein or lipoprotein barrier rich in sulfhydryl groups and disulfide bonds. These interact constantly so that a dynamic equilibrium is present. The introduction of AVP is followed by the interaction of the S-S bridge of the AVP with an SH group in the receptor. A new S-S bridge is formed between hormone and receptor and a new SH group on the hormone appears which reacts in turn with an S-S bridge in the receptor forming a second S-S bridge between hormone and receptor, and liberating a new SH group on the receptor. The entire process is reversible so that the newly formed receptor SH can either interact with another AVP molecule or reversibly release the first. At any given time the number of AVP molecules bound to the tissue will depend upon the AVP concentration. Sulfhydryl reagents, such as NEM, will prevent the action of AVP by interacting with free SH groups, or prevent the reverse reaction by binding the new receptor SH groups. The reversibility of a number of other disulfide interchange reactions has been reported.[17]

More difficult to explain are the observations that NEM or MMB, added to the mucosal solution at concentrations of 10^{-3} M completely inhibit AVP action whereas at lower concentration they mimic its effect. At present the data can be interpreted best on the assumption that the structure of this barrier is in a state of dynamic equilibrium with disulfide interchange reactions going on constantly, so that at any given moment a small number of free sulfhydryl groups are available. The addition of a sufficient concentration of NEM will completely inhibit this process, whereas a lower concentration in some manner catalyses the breakdown of an S-S bridge. Such a proposal is not novel in the field of sulfhydryl chemistry. A number of chemical studies of sulfhydryl and disulfide reactions have shown that the type of reaction or interaction observed is critically dependent upon the concentration of reagents employed.[18, 19]

One set of observations which seem to indicate that sulfhydryl or disulfide groups are not involved in the hormone action are those which show that iodo-

acetamide does not have a very profound influence upon hormonal effect. Similar unexplained situations have been observed in sulfhydryl chemistry before however. An SH group of a particular protein may react readily with one sulfhydryl reagent and sluggishly or not at all with another.[18, 19] Steric factors, charge, and the presence or absence of RS^- groups have been invoked to explain these "anomalous" results.

Another fact which does not fit readily into the present scheme is the reversion of the bladder to its original state of impermeability observed when the serosal solution is acidified after AVP has increased bladder permeability. No adequate explanation is apparent but further experiments with tritiated AVP are in progress and may help in elucidating the nature of the pH effect.

Although the change in water permeability, induced by the addition of NEM and other agents to the mucosal surface of the bladder, is similar to that observed following the addition of AVP to the serosal surface, there are certain differences in their action. The most striking difference is that the effect of the hormone is reversible whereas that of NEM is not. Also, when the hormone is added in the presence of sodium ion, there is an increase in oxygen consumption, but a decrease or no change is observed after NEM, suggesting that NEM may change the permeability to water but not augment sodium transport. Hence the hormone may have two separate actions, one on water permeability and the second on sodium transport. The site of these two functions may be spacially and functionally independent as has been suggested by the work of Crabbe who found that a steroid hormone, aldosterone, when added to the serosal surface of the toad bladder *in vitro*,[20] will increase sodium transport without changing water permeability.

That a protein is an integral part of the mucosal barrier of the cell is suggested by our observation that certain proteolytic enzymes (trypsin and pepsin), when placed in the solution bathing the mucosal surface cause the bladder to become more permeable. No data are presently available as to the protein involved; nor is there any indication as to how the postulated disulfide interchange reaction can bring about a change in membrane permeability.

Because the results obtained with oxytocin and vasotocin were so similar to those noted with vasopressin it may be suggested that one of the primary interactions between a neurohypophyseal hormone and its receptor site involves a disulfide interchange. The affinity of any particular hormone for a specific receptor will be determined by the molecular complementarity between the two.

Summary.—The well-known ability of arginine vasopressin (AVP) and other neurohypophyseal hormones to increase reversibly the permeability of the isolated amphibian bladder to water has been used as a measure of hormone action. It has been possible to show that this permeability change is not directly dependent upon the oxidative metabolism of the cell. Furthermore, a number of sulfhydryl reagents have been shown either to inhibit the effect of AVP or to prevent the reversibility of the permeability change. The hormone is only effective if added to the serosal surface of the bladder, although it is thought that the permeability barrier of the bladder is at or near the mucosal surface of its epithelial cells. The sulfhydryl reagents, however, have been found to inhibit AVP action when added to either surface. If added in critical concentrations (10^{-4} to 10^{-5}) to the mucosal surface, sulfhydryl reagents induce a change in permeability which is qualitatively

similar to the effect of AVP (added to the serosal surface) but is irreversible. From these observations it has been concluded that one of the primary interactions between arginine vasopressin and its receptor site in the bladder involves a disulfide-sulfhydryl interchange reaction.

We wish to thank Professors Lyman C. Craig and William H. Stein for many stimulating discussions during the course of this work and for their critical reading of this manuscript. We are also indebted to Dr. Alexander Leaf for introducing us to the amphibian bladder as a research tool and for freely discussing many of his results before publication. We are grateful to Professor V. du Vigneaud for his generous gift of the sample of vasotocin used in this study.

* Supported by a grant (U-1021) from the Health Research Council of New York City, and the Atomic Energy Commission.

† Investigator of the Health Research Council of New York City.

[1] du Vigneaud, V., *The Harvey Lectures*, **49**, 1 (1954–55).

[2] Ussing, H. H., and K. Zerahan, *Acta Physiol. Scand.*, **23**, 110 (1951).

[3] Koeford-Johnsen, V., and H. H. Ussing, *Acta Physiol. Scand.*, **28**, 60 (1953).

[4] Andersen, B., and H. H. Ussing, *Acta Physiol. Scand.*, **39**, 228 (1957).

[5] Leaf, A., J. Anderson, and L. B. Page, *J. Gen. Physiol.*, **41**, 657 (1958).

[6] Maffly, R. H., R. M. Hays, E. Lamdin, and A. Leaf, *J. Clin. Invest.*, **39**, 630 (1960).

[7] Bentley, P. J., *J. Endocrinol.*, **17**, 201 (1958).

[8] *Ibid.*, **18**, 327 (1959).

[9] Ewer, R. F., *J. Exptl. Biol.*, **29**, 173 (1952).

[10] Sawyer, W. H., and R. M. Schisgall, *Am. J. Physiol.*, **187**, 312 (1956).

[11] Sawyer, W. H., *Endocrinol.*, **66**, 112 (1960).

[12] Leaf, A., and E. Dempsey, *J. Biol. Chem.*, (in press).

[13] Rothen, A., B. F. Chow, R. O. Greep, and H. B. van Dyke, *Cold Spring Harbor Symposia Quant. Biol.*, **9**, 272 (1941).

[14] Schwartz, I. L., C. T. O. Fong, E. A. Popenoe, L. Silver, and M. A. Schoessler, *J. Clin. Invest.*, **38**, 1041 (1959).

[15] Fong, C. T. O., I. L. Schwartz, E. A. Popenoe, L. Silver, and M. A. Schoessler, *J. Am. Chem. Soc.*, **81**, 2592 (1959).

[16] Leaf, A., *J. Cell. Comp. Physiol.*, **54**, 103 (1959).

[17] Jensen, E., *Science*, **130**, 1319 (1959).

[18] Cecil, R., and J. R. McPhee, *Ad. Prot. Chem.*, **14**, 255 (1959).

[19] Pihl, A., and L. Eldjarn, *Proc. Fourth Int. Cong. Biochem.*, **13**, 43 (1959).

[20] Crabbe, J., *Clin. Res.*, **8**, 227 (1960).

*RELATION OF CHEMICAL ATTACHMENT TO PHYSIOLOGICAL ACTION OF VASOPRESSIN**

By Irving L. Schwartz, Howard Rasmussen,† Mary Anne Schoessler, Lawrence Silver, and Conrad T. O. Fong

THE MEDICAL RESEARCH CENTER OF BROOKHAVEN NATIONAL LABORATORY, THE ROCKEFELLER INSTITUTE, AND THE NEW YORK UNIVERSITY COLLEGE OF MEDICINE

Communicated by Donald D. Van Slyke, August 5, 1960

We have suggested in preliminary publications[1, 2] that the antidiuretic action of vasopressin in the mammalian kidney may depend on a sulfhydryl-disulfide interchange involving SH groups on the receptor and the S–S bridge of the hormone. Strong support for this concept is presented in the preceding paper[3] where it is shown that the vasopressin-induced increase in the permeability of a simple living membrane (the toad urinary bladder) depends on the availability of free and accessible SH groups on the membrane and where it is also shown that this action of vasopressin does not depend directly on the oxidative metabolism nor on the sodium transporting function of the membrane.

The present work, employing tritium-labeled arginine vasopressin, provides direct evidence for this concept, for it has been found that there is a close correlation between the formation of hormone-receptor disulfide bridges and the physiological action of vasopressin on membrane permeability.

Experimental Methods.—Highly purified arginine vasopressin was prepared from a partially purified Pitressin powder[4] by ion exchange chromatography[5] and then tritiated by exposure to H^3 gas under electric discharge.[6] Easily exchangeable tritium was removed by repeated solution and lyophilization from dilute acetic acid, following which the tritiated arginine vasopressin (H^3AVP)‡ was separated from radiation degradation products by the same chromatographic procedure[5] used originally to purify the natural hormone. This work is described in detail by Fong, Silver, Christman, and Schwartz.[7]

In addition to chromatographic and electrophoretic evidence for purity, rat antidiuretic assay and rat pressor assay showed biological activity of approximately 400 pressor units/mg, which is comparable to potencies that have been reported for highly purified preparations of natural arginine vasopressin.[8–12] The specific radioactivity of the purified H^3AVP was 426 μc/mg.

The procedure used to measure the hormonal effect on the permeability of the bladder is described in the preceding paper.[3] In our hands Pitressin, Pitocin, Syntocinon,[13] arginine vasopressin, H^3-arginine vasopressin, lysine vasopressin and arginine vasotocin[14] all showed activity of the same relative potencies as reported by Sawyer.[15]

Exposure of bladders to tritiated arginine vasopressin (H^3AVP) was carried out in the same manner as experiments with unlabeled hormone or by a modified procedure in which the bladders were filled with Ringer's solution diluted 1:5, and then tied as balloons as they were slipped off a tapered glass tube. These "balloons" were then placed in an aerated bath of full-strength Ringer's solution. Hormone and/or sulfhydryl-binding reagents were added to the latter solution to give final concentrations in the bath as follows: H^3AVP $- 2.5 \times 10^{-8}$ M (ca.

0.025 μg/ml) (ca. 10 mμ/ml); N-ethylmaleimide, 10^{-3} to 10^{-5} M; *p*-chloromercuribenzoate, 10^{-3} to 10^{-5} M; methyl mercuric bromide, 10^{-3} to 10^{-6} M. In most of the experiments in which sulfhydryl-binding reagents were used, the bladders were incubated in Ringer's solution containing these compounds for a period of 20 min and then exposed to H^3AVP for 30 min. At the same time control bladders were incubated for 20 min in plain Ringer's solution and then exposed to H^3AVP for 30 min. In all experiments involving H^3AVP the bladders were removed from the bath, cut off below the tie to eliminate the possibility of measuring hormone which had entered the cells through the cut edges, and then placed either in ethanol or boiling water for 2 to 3 min to destroy enzymatic activity. In either case the ethanol or the water contained N-ethylmaleimide, 10^{-3} M, in order to inhibit thiol-disulfide interchange reactions which might take place during the course of the denaturation of the bladder protein.

The bladders were then treated to remove adsorbed and/or electrovalently bound radioactivity, that is, hormone attached to tissue through ionic and hydrogen bonds. In the early experiments this was accomplished by electrodialysis for 48 hr followed in some instances by treatment with 8 M urea. Bladders were placed in viscose tubing which was tied at both ends and immersed in a large container filled with distilled water at 10°C. Pt electrodes were placed in the container and a potential difference of 800 to 1,000 volts was maintained throughout the time of dialysis. The water was changed repeatedly. In later experiments, hormone attached to the bladder by noncovalent forces was removed by subjecting the tissue to a series of successive immersions (for 2–3 min with shaking) in 1.0 M NaCl, 0.15 M $NaHCO_3$ (pH 8.5), and Na Acetate buffer (pH 4.0). Bladders subjected to this latter procedure showed the same level of residual tightly bound radioactivity as those given the more lengthy electrodialysis (Table 1). In some experiments bladders were given an additional wash in a solution of 8 M urea. This treatment caused no change in the level of residual radioactivity. After removal of adsorbed and/or electrovalently bound hormone, some of the bladders were subjected to conditions specific for the cleavage of disulfide bonds, that is, each bladder in this group was placed in a vial containing 3.0 ml of 0.1 M cysteine in 0.15 M $NaHCO_3$ (pH 8.5). The remaining bladders served as controls. Each bladder in the control group was placed in 3.0 ml of 0.1 M glycine in 0.15 M $NaHCO_3$. The vials were gassed with nitrogen, closed tightly with screw caps, and incubated in a water bath at 37°C with shaking for 8 hr. After incubation the bladders were removed from the vials and blotted on gauze. Each bladder was weighed to the nearest mg and placed in a low-potassium, low-background glass vial[16] containing 2.5 ml of a solution of Hyamine 10-X[17] in methanol, which dissolved the tissue. Solution of the tissue could be hastened by placing the closed vials in a water bath maintained at 40–45°C. When the bladders were completely dissolved, 12 ml of toluene containing 600 mg per cent diphenyloxazole (phosphor) and 10 mg per cent bis-phenyloxazolylbenzene (spectrum shifter) were added to each vial. The samples were chilled to −10°C, and tritium activity was assayed in a Packard Tri-Carb liquid scintillation counter. Quenching due to Hyamine, methanol, and tissue was measured by addition of an internal standard (H^3-Butanol) as described by Vaughan, Steinberg, and Logan.[18] Results were calculated as counts per min, corrected for quenching, per gram bladder tissue.

Results.—Effect of AVP and H^3AVP on membrane permeability: Following addition of H^3AVP to the serosal bath the change in water permeability of the isolated bladder was similar in all respects to that observed after the same dose of nonlabeled AVP (Table 2, compare columns 1 and 3). This indicates that substitution of tritium for hydrogen atoms in the vasopressin molecule did not decrease hormonal potency or alter in any way the action of the hormone on the permeability to water of the bladder.

Binding of H^3AVP to bladder tissue: Following labeling of the bladders by addition of H^3AVP to the serosal bath, inactivation of enzymes with alcohol or heat and removal of loosely bound (adsorbed, salt-linked, hydrogen-bonded) label, significant amounts of radioactivity still remained associated with the bladder tissue (Table 1). From 50 to 80 per cent of this residual tightly bound radioactivity was removed by incubation of the bladders under mild conditions (0.1 *M* cysteine, pH 8.0–8.5) which reduce disulfide bonds but do not hydrolyze peptide bonds (Table 1, Fig. 1*A*). This indicates that the loss of radioactivity from the bladder occurred because an S–S bond between the bladder and the hormone had been split.

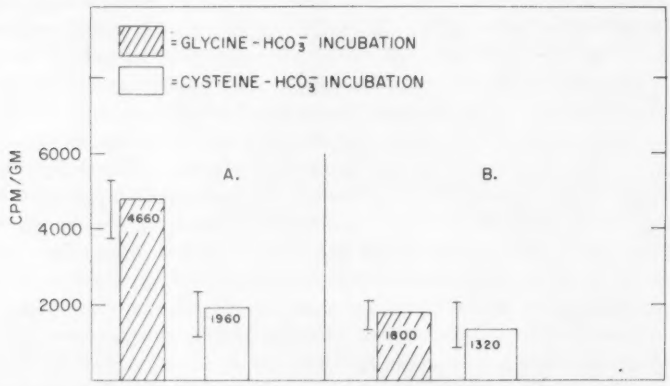


FIG. 1.—Radioactivity of H^3AVP labeled bladders after treatment with cysteine to remove disulfide bound H^3 as compared with the radioactivity of H^3AVP labeled bladders treated with glycine as controls. The vertical lines to the left of the bars give the range of values derived from six experiments. (*A*) Bladders were removed from the bath after 30 min of exposure to H^3AVP when hormone action was near maximal. (*B*) Bladders were exposed to H^3AVP for 30 min and then washed with fresh Ringer's solution for a period of 45 min at the end of which time hormone action was no longer manifest, that is, the bladder wall had regained its original level of impermeability.

When bladders were acted upon by tritiated vasopressin and then washed in fresh Ringer's solution several times during a 45-min period, the natural resting impermeability of the bladder wall was regained. When bound radioactivity was measured after this period of washing, no cysteine-releasable radioactivity could be demonstrated (Fig. 1*B*).

Effect of SH-binding reagents: As noted in the preceding paper[3] when N-ethylmaleimide (NEM), at a final concentration of 10^{-3} Molar, and other SH-binding reagents were added to the bath in which the bladders were immersed, the action

TABLE 1

RADIOACTIVITY OF BLADDERS LABELED WITH H^3AVP AND TREATED SUCCESSIVELY TO REMOVE NONCOVALENTLY BOUND H^3 AND DISULFIDE-BOUND H^3*

Conditions for removal of loosely bound H^3	Radioactivity Bound to Bladder Tissue: After glycine incubation (cpm/gm tissue)	After cysteine incubation (cpm/gm tissue)
Alcohol-NEM	5,380	2,100
1 *M* NaCl–pH 4–pH 8	3,900	2,340
	4,950	2,380
	5,530	2,550
Boiling H_2O-NEM	4,600	1,980
1 *M* NaCl–pH 4–pH 8	4,320	1,550
	5,080	1,250
Boiling H_2O-NEM	4,420	1,800
Electrodialysis	3,323	1,010
	Mean = 4611	Mean = 1884

* The difference in the level of labeling after incubation with cysteine as compared with the level of labeling after the control incubation with glycine is a minimum measure of the radioactivity bound to the bladder through disulfide linkage.

TABLE 2

EFFECT OF AVP, H^3AVP, AND N-ETHYL MALEIMIDE (NEM) ON WATER MOVEMENT

Period (20 Min)	AVP Weight loss (μgm/min/cm²) 1	2	H^3AVP Weight loss (μgm/min/cm²) 3	4
	Bladders filled with 1/5 Ringers; immersed in normal Ringers		Bladders filled with 1/5 Ringers; immersed in normal Ringers	
1	60	50	60	100
2	30	40	120	80
	No change	Change serosal bath to Ringers containing 10^{-3} *M* NEM	No change	Change serosal bath to Ringers containing 10^{-3} *M* NEM
3	80	190	80	360
4	60	160	80	480
	AVP added (0.025 μgm/ml of serosal bath)		H^3AVP added (0.025 μgm/ml of serosal bath)	
5	1,760	180	1,960	410
6	1,940	130	2,380	340
7	1,850	170	1,710	330

of AVP on water permeability was completely inhibited (Table 2), although a variety of nonthiol-binding metabolic inhibitors (dinitrophenol 10^{-3} *M*, Na fluoroacetate 10^{-2} *M*, K cyanide 10^{-3} *M*, Na Fluoride 10^{-2} *M*, ouabain 10^{-4} *M*) did not block hormone action. It is noteworthy that in the experiments shown in Table 2 NEM itself produced a small increase in the permeability to water of the bladder wall, but this increase was not comparable qualitatively or quantitatively to that produced by the action of AVP or H^3AVP on unblocked bladders. In parallel experiments in which bladders were pretreated with SH binding reagents before exposure to H^3AVP, there was a striking reduction in the amount of radioactivity that remained attached to the bladder tissue after the procedures designed to remove noncovalently bound tritiated hormone (Fig. 2). This reduction in bladder labeling was referable primarily to the absence of the normally evident cysteine-releasable component and only slightly to a reduction of the component that was resistant to treatment with cysteine (Fig. 2). These findings indicate that in all probability the labeled hormone could no longer form disulfide bridges with blocked SH groups on receptor elements of the bladder and, therefore, support the hypothesis that an accessible SH group on the target organ interacts with the hormone disulfide bridge as a necessary step in the physiological action of vasopressin.

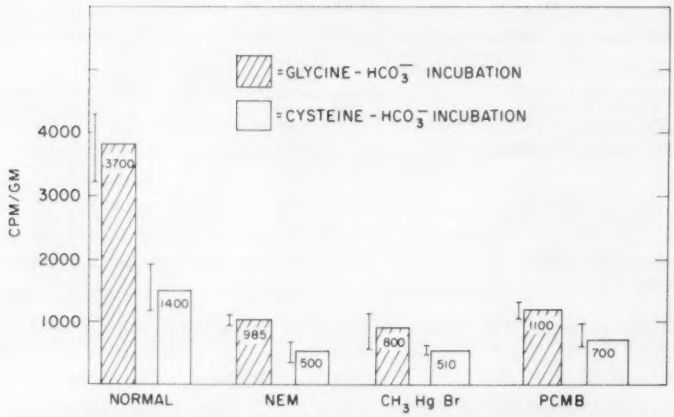


FIG. 2.—Effect of prior exposure to sulfhydryl-binding reagents on the labeling of bladders subsequently treated with H^3AVP. (NEM = N-ethylmaleimide, CH_3HgBr = methyl mercuric bromide, PCMB = *p*-Chloromercuribenzoate.) The vertical lines to the left of the bars give the range of values derived from eight control experiments and six experiments with each SH reagent.

In one experiment the serosal surfaces of the bladders were treated with methyl mercuric bromide (CH_3HgBr), 10^{-3} Molar, for 10 min, and then with cysteine, 10^{-1} Molar, for 10 min, before exposure to H^3AVP. It was found that both the hormonal activity and the binding of hormone to bladders through disulfide linkage showed the same correlation as observed with bladders receiving no pretreatment prior to exposure to H^3AVP. These data indicate that the binding of CH_3HgBr by the receptor SH group can be reversed with the resultant regeneration of a physiologically responsive receptor site.

Effect of pH: Bentley[19] found that bubbling a gas mixture containing 5% CO_2 – 95% O_2 through the serosal bath resulted in a fall in pH to 6.5. Although hormone action was almost completely inhibited under these circumstances, a systematic study of the relation of this inhibition to the pH of the bath was not reported. For this reason we carried out studies of the effect of graded changes in the pH of the serosal solution upon the physiological action as well as the binding of arginine vasopressin. It was found that when the serosal bath was kept at pH 7.0 or below, the hormone action was much less than at pH 7.4 (Fig. 3). Parallel experiments showed that, when the action of H^3AVP on membrane permeability was blocked by lowering of pH of the serosal bath, the disulfide-bound component was no longer demonstrable (Fig. 4). These results constitute additional evidence that the physiological action of vasopressin involves a sulfhydryl-disulfide interchange since, in this reaction, it is the mercaptide ion that attacks the disulfide bond. At a pH of 7 or lower, the ionization of the sulfhydryl group is diminished and therefore conditions are unfavorable for the interchange reaction to proceed.

Binding of H^3AVP in choline Ringer's solution: As noted previously[3] when the bladder is immersed in choline Ringer's solution instead of sodium Ringer's solution, the response to a standard dose of AVP was decreased by 50 per cent. However, after exposure of bladders to H^3AVP, the amount of cysteine-releasable as

well as totally bound radioactivity proved to be the same in either solution (Fig. 5). Thus, the substitution of choline for sodium in the Ringer's solutions diminishes the effectiveness of AVP in some way other than by altering the binding of hormone to receptor. It seems likely that in some more or less complex fashion the proper functioning of the sodium pump conditions the bulk flow of water through the tissue.

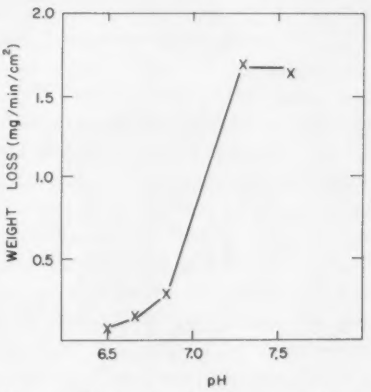


FIG. 3.—Effect of pH in the serosal bath on the response of the bladder to AVP.

Discussion.—The data presented in this and the preceding paper are consistent with the hypothesis that a primary interaction between arginine vasopressin and its receptor site in the amphibian urinary bladder involves an interchange between the disulfide bond of the hormone and a sulfhydryl group on the receptor.

The conclusion that vasopressin is bound to its receptor through a disulfide bridge is based on the fact that treatment with thiols, in particular l-cysteine, resulted in dissociation of labeled hormone from its attachment to bladder tissue. Splitting of disulfide bonds by thiols is a well established reaction which is fully described and documented in recent reviews by Cecil and McPhee[20] and Boyer.[21] The cleavage may occur by actual reduction of the original S–S bond, as in

$$R—S—S—R + 2R'SH \rightarrow R'—S—S—R' + 2RSH,$$

or by an exchange reaction with formation of a mixed disulfide, as in

$$R—S—S—R + R'SH \rightarrow R—S—S—R' + RSH,$$

and, experimentally, several reductive and interchange components can be identified from the reaction mixture. At neutral and alkaline pH the reaction is initiated by the mercaptide ion. The important point is that the net result is the splitting of the original S–S bond.

Many reports of the thiol reduction of protein disulfides indicate that the mild conditions usually employed are specific in that disulfide bonds are cleaved without concomitant hydrolysis of peptide bonds.[20, 21] Pechere, Dixon, Maybury, and Neurath,[22] in particular, demonstrated the reliability and specificity of alkaline cysteine and alkaline thioglycollate as a general laboratory procedure for reduction of S–S bonds. Thus, there is considerable evidence to support our conclusion that the release of radioactivity from the cysteine-treated bladder is due to cleavage of an S–S bond between hormone and bladder.

The fact that treatment with cysteine does not release all of the bound radioactivity suggests that either some of the hormone-bladder disulfide bonds are buried and less accessible to attack as a result of the combination, or that some hormonal sulfur is associated with the bladder tissue other than through disulfide linkage. It is also possible that the cysteine-resistant component of bladder-bound

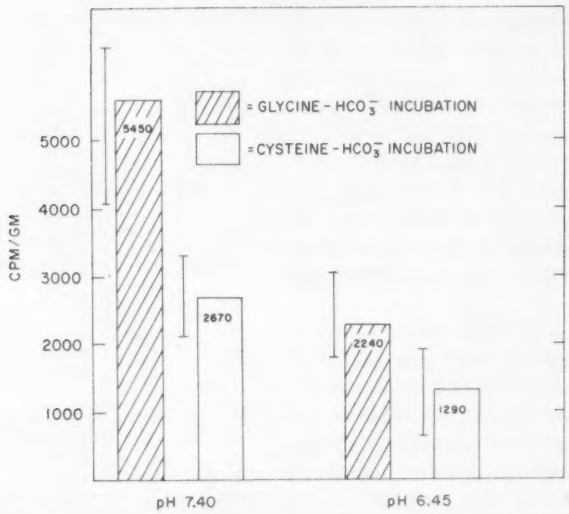


FIG. 4.—Effect of pH in the serosal bath on the binding of H^3AVP to the bladders. The difference between the cross-hatched and open bars is a measure of the disulfide bound radioactivity released from the bladders as a result of treatment with cysteine.

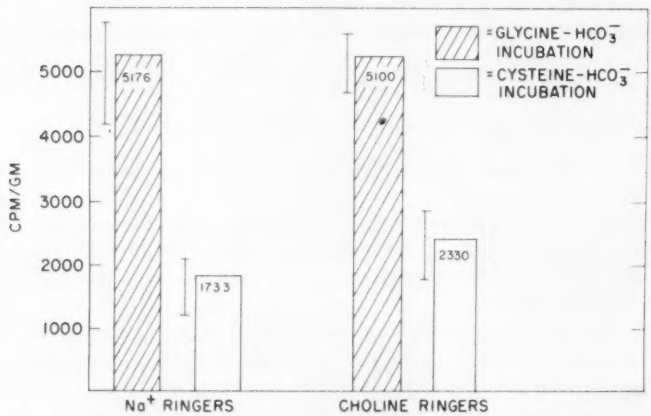


FIG. 5.—Binding of H^3AVP to bladders in sodium and in choline Ringer's solutions.

radioactivity represents hormone linked to bladder through nonsulfur centers in the vasopressin molecule.

Under normal conditions, the cysteine-releasable (i.e., the disulfide-linked) component of bound radioactivity is larger than the cysteine-resistant one; but when the physiological action of the hormone is terminated or blocked, the cysteine-releasable component is minimal or absent in every instance, whereas the cysteine-resistant component is of a similar order of magnitude (Figs. 1, 2, and 4). This

suggests strongly that the cysteine-resistant bound radioactivity is a result of hormone-bladder interactions that are without functional significance.

Of interest in the above context was the observation that, under proper conditions, the inhibitory effect of methyl mercuric bromide (CH_3HgBr) upon both hormonal action and cysteine-releasable hormonal binding could be reversed by washing the CH_3HgBr out of the bladder tissue with an excess of cysteine.

Little information is available concerning the nature of the processes by which the bladder regains its impermeability when the serosal bath containing hormone is replaced with fresh Ringer's fluid. It is possible that dissociation of hormone from receptor may be a simple reversal of the original combination or is effected by a disulfide reductase.[23] In either case definitive inactivation of hormone may then involve dimerization,[24] formation of mercapto derivatives and attack by peptidases.[25, 26] Alternatively, hormone removal may be initiated by intracellular mercaptides as suggested in the preceding paper, or other nonenzymatic reducing agents which may accumulate in the region of the receptor sites. Accordingly, it is of interest that addition of cysteine to the serosal bath (final concentration 0.01 to 0.1 *M*) increases the rate at which impermeability is re-established[27] and, presumably, therefore, at which hormone is split off from the receptors. Probably the most convincing indication that there is reductive cleavage of hormone from receptor is the fact that disappearance of physiological activity is associated with disappearance of disulfide-bound radioactivity (Fig. 1).

The hypothesis that a thiol-disulfide interchange involving receptor and hormone is concerned critically in the mechanism of action of vasopressin is consistent with the dramatic inhibition of hormone action observed when the pH of the serosal bath is lowered below 7. Work with simple systems[28] has demonstrated that the nucleophilic attack of a mercaptide ion on an S–S bond will initiate a sulfhydryl-disulfide exchange reaction. As the pH of the solution bathing the membrane drops below neutrality, receptor mercaptide is more strongly protonated and less able to react with the hormone S–S bridge. As detailed under *Results* and illustrated in Figures 3 and 4, there is a striking parallelism between the inhibition of hormone action and the inhibition of hormone-receptor disulfide-bonding as the pH of the serosal bath is lowered.

If we are correct in hypothesizing that an exchange reaction between hormone S–S and receptor SH is required for the action of vasopressin on membrane permeability, this explains previous reports in the literature which indicate that the integrity of the S–S bridge of the posterior pituitary hormones is important for their biological action. Van Dyke, Chow, Greep, and Rothen[29] in 1942 reported that reductive treatment with cysteine and thioglycollate caused 50 to 100 per cent loss, respectively, of the biological activity of their neurohypophyseal protein. This finding has been confirmed both for Pitressin[23, 26] and purified AVP.[27] Earlier, Sealock and du Vigneaud[30] had shown that reduction of Pitressin with cysteine (followed by benzylation to prevent reformation of the S–S bridge) resulted in complete inactivation of the hormone with regard to its pressor effects. Oxytocin also is inactivated by reduction, as shown by Audrain and Clauser.[31] These workers assayed the hormone under anaerobic conditions, thus preventing reoxidation and closure of the disulfide bridge without necessitating the additional alteration of the molecule ensuing from preparation of an aryl or alkyl derivative. Re-

duction with cysteine caused profound inactivation, but on reoxidation approximately 90 per cent of the original activity was restored.

It has been tempting to speculate on how an SH, S–S interchange between hormone and receptor could increase the water permeability of the bladder. A number of models can be considered. For example the hormone, by interacting with SH and S-S groups of the receptor, may induce profound conformational changes in the protein of the diffusion barrier by disrupting critical disulfide cross-linking (Fig. 6) and thereby opening channels through which water, urea, and Na^+ can move. The nature of the groups lining these channels could determine the selectivity of this passive process. Alternatively, a hormone-induced conformational change may modify an ordered water lattice around these structural

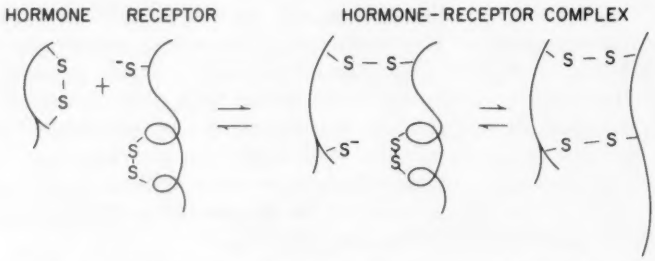


FIG. 6.—A hypothetical mechanism whereby combination of vasopressin with a membrane brings about a conformational change in a receptor protein.

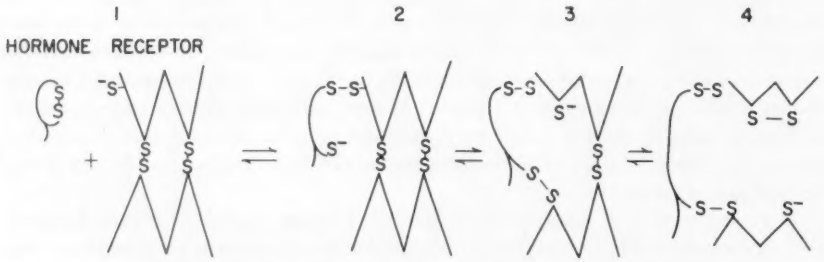


FIG. 7.—A hypothetical model illustrating linear propagation of a series of SH, S-S interchanges producing a separation of fibrillar elements in a protein diffusion barrier.

proteins,[32] i.e., melt the "ice" surrounding them, and thereby allow penetration of water and small molecules through this critical layer. Another possibility is that the hormone induces a separation of disulfide-linked fibrillar elements (Fig. 7). This scheme does not require identity of the receptor element with the critical component of the diffusion barrier, since the hormone-receptor reaction might serve only to initiate a series or wave of SH, S–S interchanges which ultimately could reach and affect the diffusion barrier even if the latter were located a full cell-length from the receptor sites. Indeed, the far reaching effects of such a biological chain reaction are consistent with the finding that vasopressin acts on the bladder only when the hormone has access to the serosal surface, whereas the

principal diffusion barrier of the bladder is at or on the mucosal surface.[33, 34] It is noteworthy that wavelike reactions of SH, S–S interchange already have been implicated in a variety of major biological phenomena, e.g., in aspects of protein denaturation, blood clotting, and the structural changes ensuing during the mitotic cycle.[28]

It is premature now to favor any one model or scheme, although it seems inescapable that, at the molecular level, the mechanism of action of vasopressin on membrane permeability involves a sulfhydryl-disulfide interaction.

Summary.—Tritium-labeled arginine vasopressin has been prepared and found to retain the full activity of highly purified natural vasopressin in increasing the permeability to water of a living membrane (the toad bladder). Using this preparation it has been possible to study the conditions under which this hormone is bound and released from the bladder and the relation of these phenomena to its physiological action. The results indicate that not only the fixation of vasopressin to its receptor but also its physiological activity depend on the presence of an accessible SH group in the bladder tissue and on the formation of a hormone-receptor disulfide bond. To our knowledge this is the first identification of a physiologically significant chemical bond formed between a hormone and a target organ. Reasons are given for suspecting that the formation of this bond, either per se or by initiating a wavelike reaction of sulfhydryl-disulfide interchange, may account for the action of vasopressin and other disulfide peptide hormones on the permeability of cell membranes.

We gratefully acknowledge the kind and helpful support of Dr. A. B. Gutman and Dr. Lewis Thomas.

The sample of arginine vasotocin used in this study was made available to us through the generosity of Dr. Vincent du Vigneaud.

* This work was supported by the U.S. Atomic Energy Commission, the Health Research Council of New York (Grant U-1021), the National Science Foundation, and the J. M. Kaplan Fund.

† Investigator of the Health Research Council of New York City.

‡ The following abbreviations are used: AVP, arginine vasopressin; LVP, lysine vasopressin; ADH, antidiuretic hormone, which is used synonymously with lysine or arginine vasopressin; AVT, arginine vasotocin; H^3AVP, tritiatedar ginine vasopressin. The following trade names are used: Pitressin—Parke Davis and Co.; Pitocin—Parke Davis and Co.; Syntocinon—Sandoz Inc.

[1] Schwartz, I. L., C. T. O. Fong, E. A. Popenoe, L. Silver, and M. A. Schoessler, *J. Clin. Invest.*, **38**, 1041 (1959).

[2] Fong, C. T. O., I. L. Schwartz, E. A. Popenoe, L. Silver, and M. A. Schoessler, *J. Am. Chem. Soc.*, **81**, 2592 (1959).

[3] Rasmussen, H., I. L. Schwartz, M. A. Schoessler, and G. Hochster, these PROCEEDINGS, **46**, 1278 (1960).

[4] Kindly supplied by Parke, Davis and Co. through the courtesy of D. A. McGinty and L. T. Blouin.

[5] Ward, D. N., and R. Guilleman, *Proc. Soc. Exp. Biol. and Med.*, **96**, 568 (1957).

[6] Dorfman, L. M., and K. E. Wilzbach, *J. Phys. Chem.*, **63**, 799 (1959).

[7] Fong, C. T. O., L. Silver, D. Christman, and I. L. Schwartz, these PROCEEDINGS, **46**, 1273 (1960).

[8] Turner, R. A., J. G. Pierce, and V. du Vigneaud, *J. Biol. Chem.*, **191**, 21 (1951).

[9] Popenoe, E. A., J. G. Pierce, V. du Vigneaud, and H. B. Van Dyke, *Proc. Soc. Exp. Biol. and Med.*, **81**, 506 (1952).

[10] Condliffe, P. G., *J. Biol. Chem.*, **216**, 455 (1955).
[11] Katsoyannis, P. G., D. T. Gish, and V. duVigneaud, *J. Am. Chem. Soc.*, **79**, 4516 (1957).
[12] du Vigneaud, V., *The Harvey Lectures* (New York: Academic Press, 1954–55), p. 1.
[13] Kindly supplied by Sandoz, Inc., through the courtesy of S. Gimpel.
[14] Katsoyannis, P. G., and V. du Vigneaud, *J. Biol. Chem.*, **233**, 1352 (1958).
[15] Sawyer, W., *Endocrinol.*, **66**, 112 (1960).
[16] Wheaton Spectrometer vials, Wheaton Glass Co., Millville, New Jersey.
[17] Para-diisobutyl cresoxy methoxy ethyl dimethyl benzyl ammonium chloride monohydrate, originally prepared according to Passman, J. M., N. S. Radin, and J. A. D. Cooper, *Anal. Chem.*, **28**, 484 (1956); later obtained commercially from the Packard Instrument Company, Inc., LaGrange, Illinois.
[18] Vaughan, M., D. Steinberg, J. Logan, *Science*, **126**, 446 (1957).
[19] Bentley, P. J., *J. Endocrin.*, **17**, 201 (1958).
[20] Cecil, R., and J. R. McPhee, *Adv. Prot. Chem.*, **14**, 255 (1959).
[21] Boyer, P., in *The Enzymes*, ed. P. D. Boyer, H. Lardy, and K. Myrback (New York: Academic Press, 1959) vol. 1, p. 511.
[22] Pechere, J., G. H. Dixon, R. H. Maybury, and H. Neurath, *Biochem. Biophys. Acta*, **31**, 259 (1959).
[23] Dicker, S. E., and A. L. Greenbaum, *J. Physiol.*, **141**, 107 (1958).
[24] Ressler, C., *Science*, **128**, 1281 (1958).
[25] Lawler, H. C., and V. du Vigneaud, *Proc. Soc. Exp. Biol. and Med.*, **84**, 114 (1953).
[26] Thorn, N. A., "Mammalian Antidiuretic Hormone," *Physiol. Rev.*, **38**, 169 (1958).
[27] Unpublished data.
[28] Jensen, E. V., *Science*, **130**, 1319 (1959).
[29] Van Dyke, H. B., B. F. Chow, R. O. Greep, and A. Rothen, *J. Pharmacol. and Exp. Therap.*, **74**, 190 (1942).
[30] Sealock, R. R., and V. du Vigneaud, *J. Biol. Chem.*, **54**, 433 (1935).
[31] Audrain, L., and H. Clauser, *Biochem. Biophys. Acta*, **30**, 191 (1958).
[32] Klotz, I., *Brookhaven Symposium in Biology*, No. 13 (1960).
[33] Leaf, A., *J. Gen. Physiol.*, **43**, 175 (1960).
[34] Maffly, R. H., R. M. Hays, E. Lambdin, and A. Leaf, *J. Clin. Invest.*, **39**, 630 (1960).

---

*NUCLEIC ACIDS AND METALS, II: TRANSITION METALS AS DETERMINANTS OF THE CONFORMATION OF RIBONUCLEIC ACIDS**,†

BY KEIICHIRO FUWA,‡ WARREN E. C. WACKER, ROBERT DRUYAN§, ANTHONY F. BARTHOLOMAY, AND BERT L. VALLEE

BIOPHYSICS RESEARCH LABORATORY, DEPARTMENT OF MEDICINE, HARVARD MEDICAL SCHOOL, AND THE PETER BENT BRIGHAM HOSPITAL, BOSTON

Communicated by Eric G. Ball, July 29, 1960

RNA isolated from a wide variety of biological sources, extending from simple microorganisms to mammals, contains significant amounts of metals. Metals of the first transition series, especially, are bound so firmly to RNA that they are not readily removed by chelating agents which are known to bind these ions quite avidly.[2] The functional significance of this observation has been uncertain.

The present studies demonstrate that metals in RNA maintain the secondary helical structure of the molecule through the formation of intramolecular bonds. The stabilization of the conformation of RNA by metals has been established through an examination of their effect on the characteristics of the phase-transition

which RNA undergoes on heating. As the temperature is raised, a phase-transition is observed which has been interpreted to reflect the melting-out of the ordered helical structure of RNA due to the disruption of intramolecular hydrogen bonds between the component nitrogenous bases. This melting curve may be measured either by an increase in absorbancy at 258 mμ or by a decrease of specific optical rotation.[3–6] We now find that metals of the first transition series of the periodic system affect a different parameter of the phase transition curve, i.e., A_{max}, the maximum increase in absorbancy attained as a result of heating. Variations in A_{max}, as a function of metal ion concentration, indicate that metal bonds, in addition to hydrogen bonds, play a decisive role in the stabilization of the molecular conformation of RNA.

Materials and Methods.—RNA was prepared by phenol extraction from beef liver,[7] rat liver,[8] and tobacco mosaic virus.[9] Beef liver RNA was successively precipitated with 2 volumes of ethanol 5 times from 0.1 versene and 3 times from 1×10^{-3} *M*, 1,10-phenanthroline to prepare "metal-free" RNA. Ammonium acetate buffer was prepared by titration with acetic acid. Both this buffer and NaCl were extracted 3 times with 0.01 per cent dithizone at acid, neutral, and alkaline pH to remove contaminating metals. Urea was recrystallized twice. Metals were added as the chloride salts; all, except magnesium, were prepared from reagent grade chemicals of known purity; magnesium chloride was prepared by dissolving magnesium metal (Johnson-Matthey Ltd., London, "Specpure") in distilled HCl and heating to dryness. Dialyses were performed in Visking tubing, prepared before use.[10] Metal-free water and acid-cleaned glassware were used throughout.[11]

RNA was dissolved in NaCl less than 48 hr before use and stored at 4°. One half to 2 hr prior to heating, this stock solution was diluted to 25 μg of RNA per ml of buffer (0.025 *M* NH_4 ($C_2H_3O_2$) in 0.025 *M* NaCl, pH 6.8); metals were added at 4°.

A Beckman DU Spectrophotometer equipped with a thermospacer and connected to an external circulator was employed for spectrophotometric measurements. Temperature was measured by a thermometer in the cell compartment. The melting curves were observed while the temperature changed at a rate of 1° per min.

Specific optical rotation was measured with a model 200S-80Q photoelectric spectropolarimeter with an oscillating prism (O. C. Rudolph and Sons) using a 10 cm water-jacketed cell; the RNA concentration was 1.25 mg/ml in 0.025 *M* NH_4 ($C_2H_3O_2$), pH 6.8, and 0.025 *M* NaCl. After the circulating water bath reached the stated temperature, ten minutes were allowed for equilibration.

Results.—The absorption of radiation by beef liver RNA, measured at 258 mμ, shows marked hypochromicity compared to that of its component bases, measured on the hydrolysate: an increase in absorbancy of 1.31 is obtained reproducibly on alkaline hydrolysis, expressed as the ratio of the absorbancy at 258 mμ after hydrolysis to the absorbancy at 10°. While ribonucleic acids from different sources have been shown to behave qualitatively similarly, quantitative differences in this hypochromic effect are characteristic of each RNA.[12]

On heating, RNA also becomes hyperchromic by asymptotically approaching a limiting value of absorbancy which is significantly less than that observed on hy-

drolysis.[12a] For native beef liver RNA there is a reproducible maximal increase of absorbancy of 1.22, expressed as the ratio of the maximal absorbancy to the absorbancy at 10°. We shall henceforth refer to this maximal increase as A_{max}.

A_{max} is markedly affected by metals of the first transition series. It can be augmented to 1.26 on removal of some of the metals intrinsic to native RNA, as described under *Methods*.[12b] The transition metal content intrinsic to native beef liver RNA provides a concentration of these metals equal to 5×10^{-6} *M* under the conditions here employed. The striking increase in A_{max} of the "metal-free" RNA is brought about by lowering the content of the intrinsic transition metals to only 1×10^{-6} *M*. Conversely, the addition of Ni^{++} ions, e.g., results in a marked *de-*

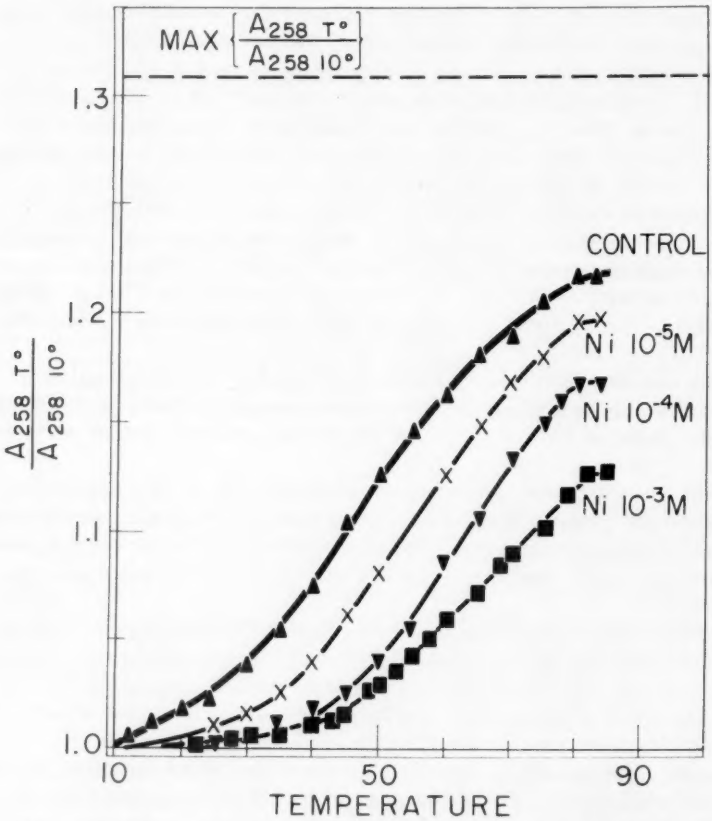


FIG. 1.—The effect of nickel on the phase transition curve of beef liver RNA. In this and subsequent figures, the increase in A_{258}, as a function of temperature is expressed as the ratio of the absorbancy at any given temperature to the absorbancy at 10°, $A_{258\,T°}/A_{258\,10°}$. The dashed line at the top indicates the maximum hyperchromicity obtained by alkaline hydrolysis. All experiments were performed under identical conditions: RNA content = 25 μgm/ml; Buffer = 0.025 *M* $NH_4(C_2H_3O_2)$ + 0.025 *M* NaCl, pH 6.8. Control of temperature as in the text. A_{max} of control (▲) = 1.22; plus 10^{-5} *M* Ni^{++} (X) = 1.20; plus 10^{-4} *M* Ni^{++} (▼) = 1.16; plus 10^{-3} *M* Ni^{++} (■) = 1.13.

pression of A_{max}, the extent of the depression being dependent on the concentration of the ion (Fig. 1). This effect, however, is not limited to nickel; in fact, it is observed with other ions of the first transition group of the periodic system shown in Table 1 which lists the hypochromic effect, induced by Cr^{+++}, Mn^{++} and Zn^{++}.

TABLE 1

EFFECT OF ADDED METALS ON A_{MAX} OF THE MELTING CURVE OF BEEF LIVER RNA

pH 6.8	A_{MAX}	pH 6.8	A_{MAX}
Control*	1.22	Li^{+} 10^{-3} M	1.22
Cr^{+++} 3×10^{-4} M	1.14	Na^{+} 10^{-1} M	1.20
Mn^{++} 3×10^{-4} M	1.17	K^{+} 10^{-3} M	1.22
Ni^{++} 10^{-3} M	1.12	Mg^{++} 10^{-3} M	1.19
Zn^{++} 10^{-3} M	1.13	Ca^{++} 10^{-3} M	1.21

* 0.025 M $NH_4(C_2H_3O_2)$ + 0.025 M NaCl, pH 68.

In the presence of 10^{-5} Cu^{++} and Fe^{+++}, precipitation occurred on heating, precluding similar observations with these metals.

In contrast neither Na^{+}, K^{+}, or Li^{+}, exemplifying the alkali metals, nor Ca^{++} and Mg^{++}, representing the alkalin eearths, in concentrations varying from 10^{-1}

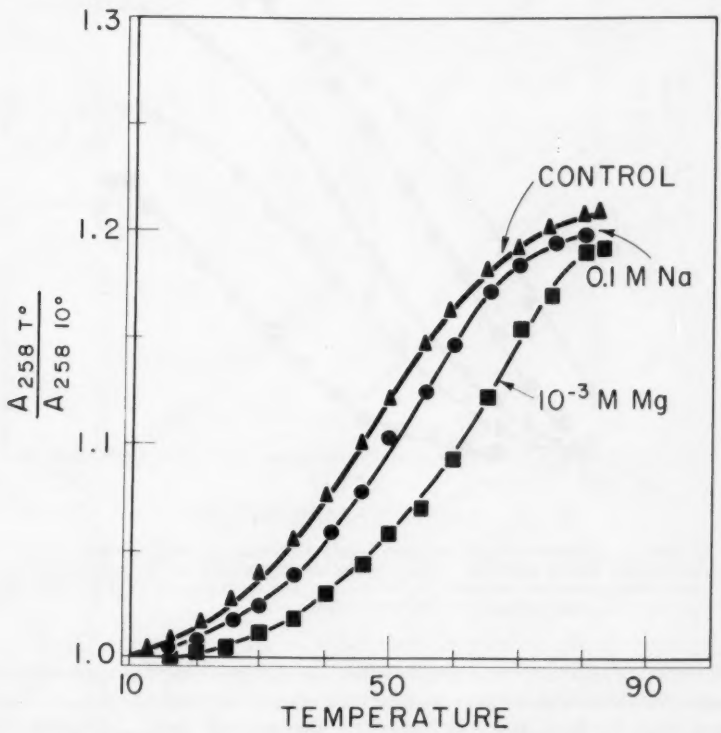


FIG. 2.—The effect of sodium and magnesium on the phase transition curve of beef liver RNA. Experimental conditions as in Figure 1. A_{max} is not altered significantly by the addition of either 10^{-1} M Na^{+} (●) or 10^{-3} M Mg^{++} (■) compared to the control (▲) even though T_m is increased.

to 10^{-3} M bring about a similar depression of A_{max} (Table 1). These elements, however, do affect T_m, the temperature at the midpoint of the phase transition curve. Results similar to those previously described have been obtained in this study as shown in Figure 2. Such changes in T_m in the presence of these elements have been interpreted in terms of hydrogen bonding which is enhanced in their presence and is diminished by urea.[3-6]

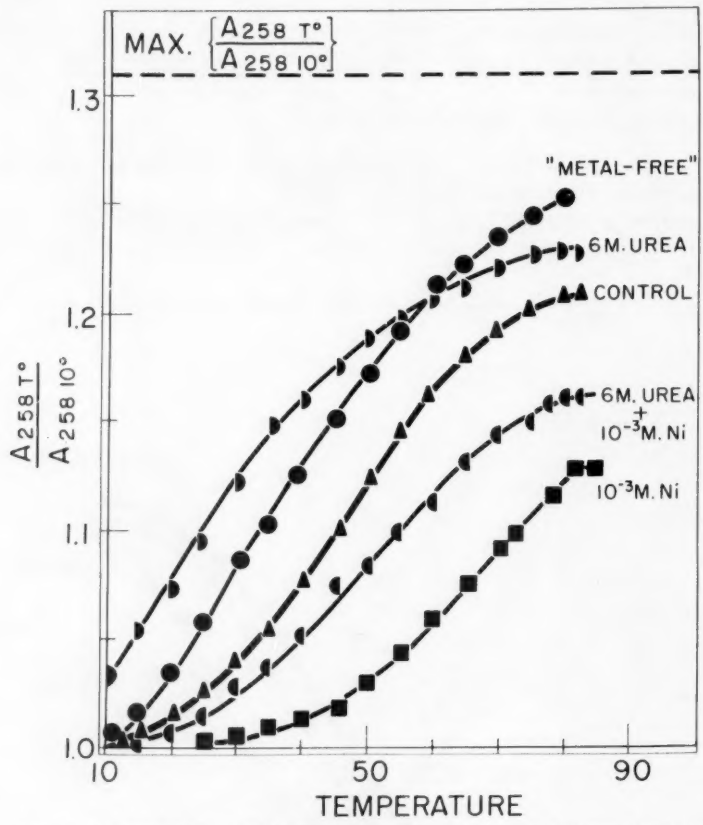


FIG. 3.—Conservation of hydrogen bonding through the establishment by Ni^{++} of tertiary folding in RNA. The experiments represented by the individual phase transition curves are indicated in the graph. The separate and joint contributions of the metal and hydrogen bonding are evident.

Parameters affecting hydrogen bonding, as expressed by T_m, appear to be independent of those which affect metal bonding, as expressed by A_{max}. This is apparent from the data shown in Figure 3. The increase in A_{max}, brought about by the removal of metals intrinsic to beef liver RNA, is here contrasted with the effect of 6 M urea which causes a marked decrease in T_m without A_{max} being affected significantly. A_{max} is, however, reduced by 10^{-3} M Ni^{++}, as previously shown in

Figure 1. In fact, an analogous depression of A_{max} by 10^{-3} M Ni^{++} is still seen, even in the presence of 6 M urea.

The specificity of metals in stabilizing the structure of RNA, and their action in this regard, as apart from hydrogen bonding, can be discerned from an experimental extension of this reasoning shown in Figure 4. When RNA is heated to the temperature at the midpoint of the phase transition curve, 49° in this instance, 50 per cent of the hydrogen bonds would be expected to be disrupted. The addition of 10^{-3} M Ni^{++}, at this juncture, instantaneously depresses the ab-

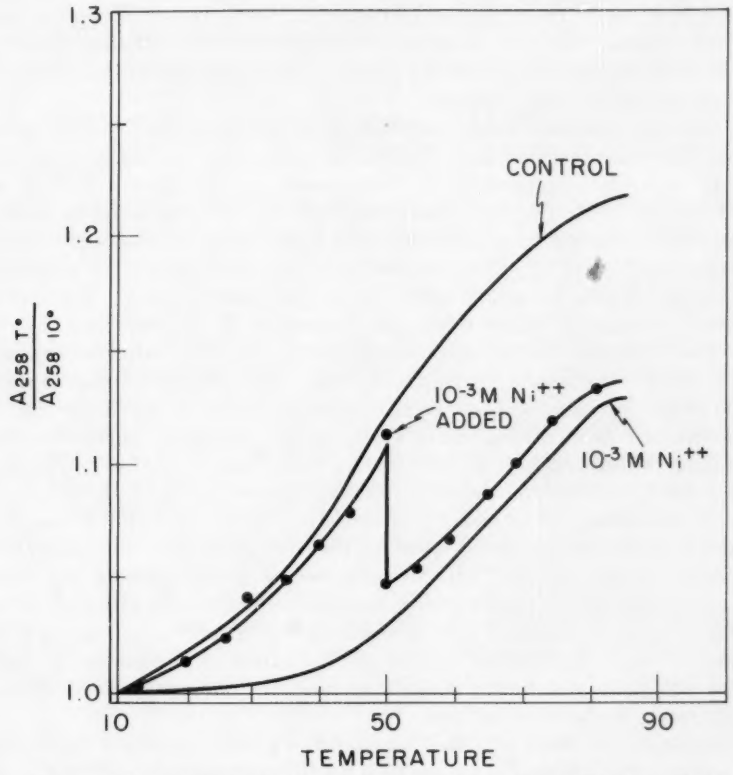


FIG. 4.—Reformation of ordered structure of native beef liver RNA at the T_m of its phase transition on addition of 10^{-3} M Ni^{++} ions. Addition of 10^{-3} M Na^{+} ions (not shown) had no effect.

sorption of radiation at 258 mμ to equal that observed in the presence of this concentration of Ni^{++} ions added prior to heating. Moreover, the phase transition eventually reaches the same value of A_{max} predicted for this concentration of Ni^{++} ions. Thus the bonds attributed to the metal can be formed even as hydrogen bonds are being destroyed.

Such findings as here presented are not limited to beef liver RNA. Analogous observations have been made with RNA from rat liver and tobacco mosaic virus.

A_{max} of TMV·RNA under the standard conditions of pH and ionic strength employed in these experiments is 1.26 while in the presence of 10^{-4} *M* Cr^{+++} it is markedly reduced to 1.07. The details of these experiments will be the subject of a separate communication.[13]

Discussion.—The existence of secondary structure in RNA has been deduced largely from the behavior of this molecule under conditions analogous to those known to affect hydrogen bonding in proteins, polymers, and copolymers.[14] On increasing the temperature, a phase transition is observed the midpoint of which is affected drastically by denaturing agents, ionic strength, solvents, and the presence of certain metals. The extinction coefficient at 258 mμ, the specific optical rotation, and the specific viscosity have served as indices of the helix-coil transition.[3-6] These experimental findings are in good agreement with expectations from theoretical treatments of such systems.[15]

Under the conditions which progressively abolish hydrogen bonding, metal ions of the first transition period are here shown to preserve and enhance the stability of the molecule as judged by the characteristics of the phase transition curves.

Metal ions of the first transition period, previously found to be constituents of native RNA prepared from many sources of widely diverging stages of evolutionary development,[2] seem to endow this molecule with the capacity to achieve intramolecular bonding by means additional to hydrogen bonding. The increase in ordered structure is apparent from the comparison of maximal chromicity of the free bases with that of all the possible constellations affecting hydrogen or metal bonding in RNA. Our results, demonstrating that metal linkages stabilize the molecule even while hydrogen bonds are being disrupted, may be viewed in the light of the numbers of conformational states available to a denatured molecule capable of resuming helical structure, as predicted by Gibbs and di Marzio.[15] The slopes of the phase transitions and the asymptotic maximum toward which each tends, observed experimentally on addition of transition metal ions to RNA, may be compared to the curves calculated, assuming the cleavage of discrete numbers of bonds of known energy content. On this basis, both a decreasing slope and maximum of absorption, as here observed with increasing concentration of transition metals, denotes increased rigidity, due—presumably—to greater preservation of helical content. Such stabilization implies the formation of intramolecular bonds in RNA which are of sufficient strength to resist the disrupting effects of increased temperature or concentrated urea.

A question now arises as to the characteristics and the number of this new species of bonds. The addition of metals to RNA in concentrations sufficient to cause a marked lowering of A_{max} neither affects the absorbancy at 258 mμ nor the specific optical rotation when measured at 10°. Thus neither the total number of internucleotide bonds nor the helical content of the molecule, dependent upon them, is increased solely by the addition of metals at low temperatures. The possibility that heating induces the formation of metal bonds is obviated by the reversibility of the melting curves on cooling. This stabilizing effect only becomes apparent on heating, when hydrogen bonds are being broken, but while persisting metal bonds attenuate the degree of the helix—coil transition. Therefore, it may be concluded that metals act to *preserve* the helical conformation of the molecule. Several interpretations may be offered to account for these observations: metal bonds may

simply be *exchanged* for existent hydrogen bonds such that the total number of bonds at low temperature remains constant, but at high temperature the metal bonds are preserved while hydrogen bonds are being broken. This view is challenged, however, by the marked effect of the metals intrinsic to RNA on A_{max}. The transition metal content of native beef liver RNA corresponds to about 1 metal atom for every 100 pairs of nucleotides.[2] After removal of some of these intrinsic metals by exposure to chelating agents, the ratio of transition metals to pairs of nucleotides is only decreased to 1:500. Despite the relatively small percentage of the total nucleotides affected directly by the intrinsic metals in native RNA and by their removal, a significant decrease in the stability of the molecule results, as evidenced by the observed increase in A_{max}. The intrinsic metals in RNA therefore exert a greater effect upon the helical structure than is accounted for on the basis of a simple, quantitative substitution of metal for hydrogen bonds.

In fact, the basis of the phenomenon appears to be different in *kind* from that implied by the above remarks. Thus, an alternative explanation for this enhanced effect of metals is required.

Characteristically, in proteins the formation of cross-linkages between peptide chains results in tertiary folding of the molecule.[16] This is exemplified most prominently by disulfide bridges[16] but also by hydrogen bonding between certain electro positive and negative groups along the chain,[17-19] hydrophobic bonds,[20] amide bonds,[21] phosphate diesters,[22] van der Waals interactions,[23] and finally by metal bridges as in trypsin[24-25] or amylase.[26] The data here reported suggest to us the existence of tertiary folding in RNA, the transition metal atoms serving as crosslinking devices, through chelation or sandwich complexes.[2] A metal bond, serving as a cross-link between nucleotide residues on adjacent turns of the helix, may be expected to fix or stabilize the hydrogen-bonded secondary structure enclosed within the tertiary loop, a well documented effect of tertiary folding in proteins.[16-27] In RNA this is recognized by the failure to achieve complete randomization of secondary structure by the rupture of hydrogen bonds on heating, i.e., a reduction of A_{max} of the melting curve, as seen in these experiments. The concept that metals stabilize RNA in this manner is strengthened by the experiments (Fig. 4) which demonstrate the ordering of the molecular structure by the addition of nickel at a temperature sufficient to have caused the disruption of half of the existent hydrogen bonds. Thus the establishment of tertiary structure allows the reformation of the helical conformation in the regions susceptible to stabilization through the creation of tertiary folds, even under conditions where the naked secondary structure is being destroyed.

That macromolecules are endowed with structural rigidity seems to be a general principle since this property is apparently required for effective biological function. Rigidity is accomplished, however, by means which seem to be adapted to the functional needs of the individual macromolecular species. Thus DNA whose function is mainly concerned with replication achieves rigidity through the two helically interwoven polynucleotide strands, held in juxtaposition by low energy hydrogen bonds.[28] The linear stability thus attained still permits the necessary separation of the two strands. In proteins, tertiary folding maintained chiefly by disulfide bridges, provides for the rigidity needed to induce the configurational specificity characteristically required in the function of enzymes, immune

bodies, and hormones.[16] It is probable that a mechanism for structural stabilization exists in RNA which promotes the specificity required for its postulated function as a template in protein synthesis.[29] The present data suggest that metals of the first transition series may be the agents responsible for this configurational property.

Since the biological activity of RNA, unlike the catalytic activity of many proteins, has not been defined in chemical terms, the role of metals in such ultimate function of RNA has not been assessed. However, metals of the first transition series stabilize the biological activity of RNA as expressed by the infectivity of TMV-RNA.[30] The constancy of metal content of TMV-RNA observed in successive analyses of different preparations eliminates the possibility that their presence is spurious.[13] The effect of their addition on A_{max} are in conformity with those observed for beef liver RNA. Studies on the implications of these compositional and conformational changes on the function of TMV-RNA are in progress.

Summary.—Certain metals of the first transition period of the periodic table are found to stabilize the ordered structure of RNA. These findings are interpreted to denote the existence of tertiary structure in RNA which is produced and maintained by metal bonds.

* This work was supported by the Howard Hughes Medical Institute and a grant-in-aid from the National Institutes of Health of the Department of Health, Education and Welfare, No. H3117 (C1), and the Lasdon Foundation.

† A preliminary report has been given (see ref. 1).

‡ Fellow of the Kettering Foundation, Dayton, Ohio.

§ Post-Doctoral Fellow of the National Institutes of Health of the Department of Health, Education and Welfare.

[1] Wacker, W. E. C., K. Fuwa, and R. Druyan, *Federation Proc.*, **19**, 316 (1960).

[2] Wacker, W. E. C., and B. L. Vallee, *J. Biol. Chem.*, **234**, 3257 (1959).

[3] Doty, P., H. Boedtker, J. R. Fresco, R. Haselkorn, and M. Litt, these PROCEEDINGS, **45**, 482 (1959).

[4] Hall, B. D., and P. Doty, *J. Molecular Biol.*, **1**, 111 (1959).

[5] Doty, P., H. Boedtker, J. R. Fresco, B. D. Hall, and R. Haselkorn, *Ann. N. Y. Acad. Sci.*, **81**, 693 (1959).

[6] Boedtker, H., *Biochim. Biophys. Acta*, **32**, 519 (1959).

[7] Kirby, K. S., *Biochim. J.*, **64**, 405 (1956).

[8] Laskov, R., E. Margolish, U. Z. Littauer, and H. Eisenberg, *Biochim. Biophys. Acta*, **33**, 247 (1959).

[9] Gierer, A., and B. Schramm, *Nature*, **177**, 702 (1956).

[10] Hughes, T. R., and I. M. Klotz, in *Methods of Biochemical Analysis*, ed. D. Glick (New York: Interscience, 1956), Vol. 3, p. 265.

[11] Thiers, R. E., in *Methods of Biochemical Analysis*, ed. D. Glick (New York: Interscience, 1957), vol. 5, p. 273.

[12] Magasanik, B., in *The Nucleic Acids*, ed. E. Chargaff and J. N. Davidson (New York: Academic Press, 1955), Vol. 1, p. 393.

[12a] In the melting curves to be shown, the phase transition has been measured by the increase in A_{259}, it has also been measured using the decrease in $[\alpha]_{546}$ as the indicator of the helix-coil transition. In comparable experiments, a correlation coefficient (r) of -0.9 between the changes in A_{258} and $[\alpha]_{546}$ has been obtained, suggesting that these parameters reflect the same physical-chemical process. Studies of viscosity also indicate a more ordered structure of RNA at increased temperatures in the presence of transition metals.

[12b] The metal content of native beef liver RNA, expressed as μgms of metal per gram RNA is as follows: Mg 500, Ca 930, Sr 26, Ba 99, Al 37, Cr 86, Mn 81, Fe 370, Ni 63, Cu 147, Zn 291.[2] After five successive precipitations from 0.1 *M* versene and three successive precipitations from 0.001 *M* 1,10-phenanthroline, the metal content is reduced: Mg 78, Ca not detected, Sr not detected, Ba 4, Al 37, Cr 31, Mn 26, Fe 155, Ni not detected, Cu not measured, Zn not detected.

[13] Wacker, W. E. C., M. P. Gordon, and B. L. Vallee, in preparation.

[14] Doty, P., *Rev. Mod. Phys.*, **31,** 107 (1959).

[15] Gibbs, J. H., and E. A. DiMarzio, *J. Chem. Phys.*, **30,** 271 (1959).

[16] Linderström-Lang, K. V., and J A. Schellman, in *The Enzymes*, ed. P. D. Boyer, A. Lardy, and K. Myrbäck, (New York: Academic Press, 1959), vol. 1, p. 443.

[17] Crammer, J. L., and A. Neuberger, *Biochem. J.*, **37,** 302 (1943).

[18] Harrington, W. F., *Biochim. Biophys. Acta*, **18,** 450 (1955).

[19] Laskowski, M., and H. A. Scheraga, *J. Am. Chem. Soc.*, **76,** 6305 (1954).

[20] Berger, A. L., and K. V. Linderström-Lang, *Arch. Biochem. Biophys.*, **69,** 106 (1957).

[21] Massey, V., and B. S. Hartley, *Biochim. Biophys. Acta*, **21,** 361 (1956).

[22] Perlman, G. E., *J. Am. Chem. Soc.*, **74,** 6308 (1952).

[23] Anfinsen, C. B., in *The Molecular Basis of Evolution* (New York: John Wiley and Sons, 1959), p. 102.

[24] Duke, J. A., M. Bier, and F. F. Nord, *Arch. Biochem. Biophys.*, **40,** 424 (1952).

[25] Gorini, L., *Biochim. Biophys. Acta*, **7,** 318 (1952)

[26] Stein, E. A., *Federation Proc.*, **16,** 254 (1957).

[27] Linderström-Lang, K. V., and J. A. Schellman, *Biochim. Biophys. Acta*, **15,** 156 (1954).

[28] Crick, F. H. C., and J. D. Watson, *Proc. Royal Soc. (London)*, **A233,** 80 (1954).

[29] Zamecnik, P. C., in *The Harvey Lectures 1958–1959* (New York: Academic Press, 1960), p. 256.

[30] Cheo, P. C., B. S. Friesen, and R. L. Sinsheimer, these PROCEEDINGS, **45,** 305 (1959).

---

*SPIN RESONANCE STUDY OF SEROTONIN-FMN INTERACTION**

BY I. ISENBERG, A. SZENT-GYÖRGYI, AND S. L. BAIRD, JR.

INSTITUTE FOR MUSCLE RESEARCH, MARINE BIOLOGICAL LABORATORY, WOODS HOLE, MASSACHUSETTS

Communicated August 15, 1960

A variety of indoles form complexes with riboflavin.[1, 2] These complexes have been interpreted as ones involving electron transfer, the indoles acting as an electron donor and the riboflavin as an acceptor. Theoretical studies[3] have shown that the positions of the energy levels of these molecules are consistent with this interpretation. The indoles have moderately high occupied energy levels, thus making them fairly good donors, while riboflavin has a low unoccupied energy level, thus making it a good electron acceptor.

This paper will report a study of serotonin and flavin mononucleotide (FMN) by the method of electron spin resonance (ESR). Tryptophan and FMN were also studied, but, for reasons stated below, not as extensively. To help identify the spectra obtained, the semiquinone of FMN at acid pH was prepared by reduction with zinc, by reduction with dithionite and photoreduction.

If an electron donor complexes with an electron acceptor, the donor giving an electron to the acceptor, the complex will not necessarily yield a signal in SR. In such a complex the donated electron will not necessarily be completely uncoupled from its partner and the two electrons will probably form a weak covalent bond

and hence the complex will yield no signal. Furthermore, even if the electrons were completely uncoupled, two free radicals in such close proximity would broaden each other's signal to a considerable extent by dipole-dipole broadening,[4, 5] and the resultant signal might be so broad as to completely avoid detection. Consequently, the technique of ESR, at least under the conditions reported here, cannot examine the complex directly. It does shed light on the complex, however, since, as will be shown, it indicates that at sufficiently acid pH the complex may dissociate into two free radicals. Ordinary charge transfer complexes, of the type extensively studied by Mulliken and his school, cannot do this. Consequently, it appears that the complexes of indoles and FMN must be classified differently. As suggested previously,[1] the complex appears to be one in which the indole gives a "complete" electron to riboflavin,† resulting in an ionic structure. It might be appropriate to call the ordinary charge transfer complexes cases of "weak charge transfer," reserving the name "strong charge transfer" for complex resulting in an ionic structure. Miller and Wynne-Jones,[6] and Bijl, Kainer, and Rose-Innes,[7] have examined complexes that also appear to be cases of strong charge transfer.

Technique.—The ESR apparatus was a Varian 4500-Spectrometer, equipped with a 12-inch magnet and a Varian 4560 100 kc modulation unit. Many features of the hyperfine structure, such as symmetry properties, are more evident when the second derivative of the absorption is obtained rather than the first. To achieve this, a dual modulation method, patterned after that of Smaller and Yasaites,[8] was adapted for use with Varian equipment. One side of the 100 kc signal was fed to the Varian 1270A low frequency phase detector. The modulation of both sweeps, but more especially, that of the low frequency sweep, was kept very low so that essentially second derivative spectra were obtained.

Solutions were examined in flat quartz cells of internal dimensions 0.1 mm × 7 mm × 52 mm. The cells were oriented in the microwave cavity so as to place the liquid in a minimum electric field.

The FMN was Nutritional Biochemical Corporation's riboflavin 5′ phosphate Sodium. The serotonin was the California Corporation for Biochemical Research Company's serotonin-creatine H_2SO_4 complex,‡ and the tryptophan was the H. M. Chemical Company's L-tryptophan. The acid used was HCl double distilled in glass. For spectra run at acid pH identical results were obtained over a wide range of acid concentration. Thus, while the results reported below were obtained mainly with undiluted constant boiling point HCl, solutions made using a few per cent HCl gave identical spectra.

The red acid semiquinone of FMN was prepared by dropping zinc into an acid FMN solution and waiting until the color of the solution became dark red. The solution was then separated from the zinc and poured into the flat cell. Reduction by dithionite instead of zinc, as well as photoreduction, gave identical results. Photoreduction was achieved by the method of Commoner and Lippincott.[9] The solution was deoxygenated and placed for 5 hr at about 10 inches from a 100-watt desk lamp. Photoreduction at neutral pH yielded a different spectrum from the photoreduction at pH < 1.

Results and Discussion.—Figure 1 shows the first derivative of the absorption for the red riboflavin free radical, while Figure 2 shows the second derivative.

Figure 2 shows mirror image symmetry with respect to the midpoint of the

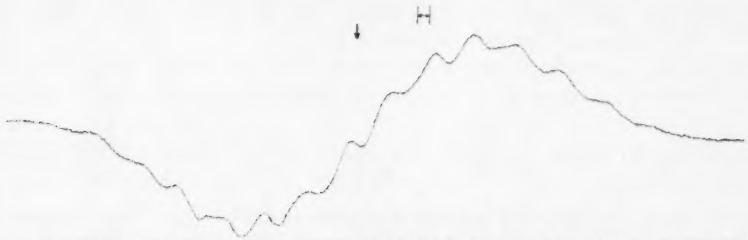


FIG. 1.—First derivative ESR spectrum of 10^{-2} *M* FMN in acid solution reduced by zinc. Marker indicates one gauss. Arrow indicates free electron resonance position.

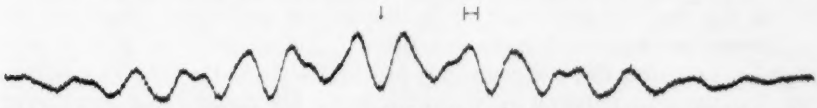


FIG. 2.—Second derivative ESR spectrum of 10^{-2} *M* FMN in acid solution reduced by zinc. Marker indicates one gauss. Arrow indicates free electron resonance position.

absorption. This is what is to be expected from a single free radical species. For two free radicals to show such marked symmetry, their *g* values must coincide. Since this is most unlikely, it seems reasonable to assume that Figures 1 and 2 are the spectra of one species of free radical and that this free radical is the riboflavin semiquinone at pH < 1.[10, 11]

Figure 1 demonstrates that the hyperfine structure shows considerable overlap. Attempts to obtain a better resolution of these lines, either by lowering the sweep modulations or by lowering the concentration of FMN, met with no success. At small modulation amplitude the lines do appear to be made up of narrower lines but the sensitivity of the instrument becomes poor under these conditions so that a complete resolution of the spectrum is lacking. Actually, a result such as this is not unexpected. In a molecule such as FMN the spin density will be large at a few places on the molecule. This will give rise to the spectrum shown. The remaining atoms will have a small but finite spin density. The nitrogen and hydrogen atoms not giving rise to the large splitting will broaden the lines causing them to overlap. The overlap thus occurs because of unresolved hyperfine structure.

Serotonin in acid gave no ESR signal and neither did FMN. Serotonin and FMN, taken together, did yield a signal.

Figure 3 shows a second derivative spectrum of serotonin and FMN. It will be seen that this spectrum is similar but not identical to the riboflavin semiquinone spectrum. It is markedly asymmetrical, indicating that there are at least two free radicals present. From the similarity to the riboflavin spectrum it may be concluded that one of these is riboflavin. The other may be serotonin. However, since the serotonin free radical spectrum is not known, it may also be a secondary free radical rather than the serotonin free radical itself. The data are therefore consistent with the following scheme.

$$\begin{array}{ccccc} S + R & \underset{k_1}{\rightleftharpoons} & S^+ R^- & \underset{k_2}{\rightleftharpoons} & \dot{S} + \dot{R} \\ & & & & \Downarrow \quad \Downarrow \\ & & & & P_1 \quad P_3 \\ & & & & \Uparrow \\ & & & & P_2 \end{array} \qquad (1)$$

FIG. 3.—Second derivative ESR spectrum of serotonin and FMN in acid. Solution was 0.16 *M* with respect to FMN and 0.59 *M* with respect to serotonin. Marker indicates one gauss. Arrow indicates free electron resonance position.

P_1 may not exist. It is included for the reasons just mentioned. P_2 and P_3 may be dimers or other products of the free radicals. The symbol S^+R^- is to be interpreted as indicating that the complex is primarily ionic and is not meant to indicate that the total charge on the complex is zero. The complex may pick up protons from the solution and hence have a net positive charge.

The above scheme may be supported by the following experiment: A solution of 10^{-2} *M* FMN in acid was placed in a quartz capillary and tested in the spin resonance apparatus. No signal was obtained. A small amount of dry serotonin was added to the sample and stirred with a fine platinum wire. This yielded a signal. Successive small increments were added and the signal became larger. Finally, however, the signal saturated and further addition of serotonin gave no further change in signal in spite of the fact that the additional serotonin went into solution. At this point, the amount of signal corresponded to one free radical per 250 FMN molecules. The interpretation may be made that the serotonin drove all of the riboflavin into the complexed form. However, the amount of free radical present is determined by the dissociation constant k_2 and this dissociation, and hence the amount of free radical in the experiment just mentioned, is not under the control of the investigator.

Tryptophan and FMN in acid also yielded a signal in ESR. However, the signal was much weaker than with serotonin, so much weaker that the hyperfine structure could not be well resolved. This finding is consistent with the observation[1] that serotonin is a stronger electron donor than tryptophan.

The existence of strong charge transfer complexes permits a unified theoretical framework in the scheme of charge transfer. At the one extreme, for weak donors and/or acceptors, there exist charge transfer complexes of the type benzene-iodine in which the ground state is overwhelmingly nonionic. For example, it is estimated that only about 3 per cent of an electronic charge is transferred from benzene to iodine.[12] At the other extreme are the well-known cases of semiquinone formation in which an electron is transferred from the reducing agent to the oxidizing agent, no stable complex being formed. As one considers donors successively stronger than benzene, one passes from weak charge transfer to cases of stronger charge transfer, finally arriving at a predominantly ionic structure. Whether this ionic structure dissociates into free radicals or not depends on whether the free energy of the free radicals, as such, is lower than that of the complex. This, in turn depends, in part, on the nature of the medium, dissociation being favored by ionizing media such as water.

No ESR signal has been observed when serotonin and FMN were mixed at neutral pH. This is evidence that the complex does not dissociate into free radicals at neutral pH. At acid pH the complex does dissociate, probably because

the riboflavin free radical is more stable in a protonated condition than in an unprotonated state. It may also be, however, that in the complex, at acid pH, the FMN picks up protons from the medium and becomes a somewhat better acceptor. There may be, therefore, a somewhat stronger charge transfer complex in strong acid.

In spite of the power of ESR techniques, a point of caution should be noted. Consider a solution of serotonin and FMN in acid. This is brown at room temperature. If the solution is frozen it becomes black, indicating, evidently, a greater degree of charge transfer complexing in the frozen state. In spite of this, such a frozen sample, when tested, yielded no ESR signal. It should not be concluded from this, however, that no free radical is present. Even if a large number of unpaired electrons were present, ESR techniques might not detect them. This may easily be shown also by freezing a 10^{-3} *M* solution of $MnCl_2$. While at room temperature a characteristic 6-line manganese spectrum is obtained, in the frozen state this signal disappears. One sees, therefore, that electron spin resonance, while a powerful tool, cannot always by itself be relied upon for determining the charge transfer state of a given system.

Summary.—Even though the technique of electron spin resonance does not measure a property of the complex directly, it does shed light on the nature of the complex. It seems most unlikely that a weak charge transfer complex could dissociate into free radicals. In the case of strong charge transfer, however, under proper environmental conditions, such dissociation seems reasonable. Spin resonance techniques yield two overlapping spectra, one of which is that of the riboflavin free radical. This is evidence that the serotonin-riboflavin complex is of the strong charge transfer or predominantly ionic type and is consistent with data previously obtained using optical absorption methods.

The authors wish to thank Dr. Bernard Smaller for advice and suggestions in establishing a dual modulation spin resonance system. They also thank Dr. Richard Bersohn for several fruitful discussions concerning the work presented here.

* This research was supported by a grant from The Commonwealth Fund, Grant No. H-2042(C3) of the National Institutes of Health, a grant from the National Science Foundation, and Grant No. 1 from the Quartermaster Research & Engineering Command, U.S. Army (Natick).

† The term "complete" is used to indicate that there is a transfer of one, or almost one, electronic charge from indole to riboflavin.

‡ The authors would like to thank Vismara Terapeutici, who supplied part of the serotonin used.

[1] Isenberg, I., and A. Szent-Györgyi, these PROCEEDINGS, **44,** 857 (1958).
[2] *Ibid.*, **45,** 1221 (1959).
[3] Pullman, B., and A. Pullman, these PROCEEDINGS, **44,** 1197 (1958).
[4] Van Vleck, J. H., *Phys. Rev.*, **74,** 1168 (1948).
[5] Weissman, S. I., *J. Chem. Phys.*, **29,** 1189 (1958).
[6] Miller, R. E., and W. F. K. Wynne-Jones, *J. Chemical Soc.*, 2375 (1959).
[7] Bijl, D., N. Kainer, and A. C. Rose-Innes, *J. Chem. Phys.*, **30,** 765 (1959).
[8] Smaller, B., and E. L. Yasaitis, *Rev. Sci. Instr.*, **24,** 991 (1953).
[9] Commoner, B., and B. Lippincott, these PROCEEDINGS, **44,** 1110 (1958).
[10] Michaelis, L., M. P. Schubert, and C. V. Smythe, *J. Biol. Chem.*, **116,** 587 (1936).
[11] H. Beinert, *J. Am. Chem. Soc.*, **78,** 5323 (1956).
[12] R. S. Mulliken, *J. Am. Chem. Soc.*, **74,** 811 (1952).

*MAMMALIAN METHYLMALONYL ISOMERASE AND VITAMIN B_{12} COENZYMES**

By Peter Lengyel, Rajarshi Mazumder, and Severo Ochoa

DEPARTMENT OF BIOCHEMISTRY, NEW YORK UNIVERSITY SCHOOL OF MEDICINE

Communicated August 10, 1960

In 1955 Flavin *et al.*[1–4] found that the main pathway of propionic acid oxidation in animal tissues proceeds via the reaction sequence propionate $\xrightarrow{\text{ATP,COA}}$ propionyl CoA[5] $\xrightarrow{\text{ATP,CO}_2}$ methylmalonyl CoA → succinyl CoA → succinate. Isomerization of methylmalonyl CoA to succinyl CoA (reaction 1) is also a key step in propionate formation by propionic acid bacteria.[6] In analogy to Lynen's β-methylcrotonyl carboxylase,[7] propionyl carboxylase (the enzyme catalyzing the carboxylation of propionyl CoA to methylmalonyl CoA) has been found to be a biotin enzyme.[8] Although reaction 1 resembled Barker's glutamate-β-methylaspartate isomerization[9] (reaction 2),[10] which requires cobamide coenzymes (vitamin B_{12} or pseudovitamin

(1) $\mathrm{COOH{-}\underset{\displaystyle CH_3}{\underset{|}{CH}}{-}COSCoA \rightleftharpoons COOH{-}CH_2{-}CH_2{-}COSCoA}$ (methylmalonyl isomerase)

(2) $\mathrm{COOH{-}\underset{\displaystyle CH_3}{\underset{|}{CH}}{-}CHNH_2{-}COOH \rightleftharpoons COOH{-}CH_2{-}CH_2{-}CHNH_2{-}COOH}$ (glutamic isomerase)

B_{12} derivatives),[11] it was not until recently realized that the Barker coenzymes are also required by methylmalonyl isomerase. This we owe to the work of Smith and Monty,[12] Gurnani *et al.*,[13] and Stadtman *et al.*[14] The first two groups of investigators used liver preparations of vitamin B_{12}-deficient rats; the latter prepared the apoenzyme from enzyme fractions of propionic acid bacteria by treatment with charcoal. Inactivation of glutamic isomerase by treatment of *Clostridium tetanomorphum* extracts with charcoal had in fact led Barker and collaborators to the discovery and isolation of the cobamide coenzymes.

In our experience, charcoal treatment of methylmalonyl isomerase preparations from sheep kidney did not affect their activity. The same was true of exposure to light or cyanide which, as shown by Barker *et al.*,[11] leads to rapid inactivation of the cobamide coenzymes. Incubation with intrinsic factor, which binds vitamin B_{12} was also ineffective. However, precipitation of the enzyme with acid from ammonium sulfate solutions, a device first used by Warburg and Christian[15] for the resolution of kidney D-amino acid oxidase into apoenzyme and its coenzyme (flavin adenine dinucleotide), led to almost complete resolution. The resolved enzyme was reactivated by dimethylbenzimidazolyl- or benzimidazolylcobamide coenzyme but adenylcobamide coenzyme, the pseudo vitamin B_{12} derivative, had little or no effect. This paper deals with these and other experiments bearing on the

strength of the apoenzyme-coenzyme binding in the native and reconstituted holoenzyme.

Materials and Methods.—The methylmalonyl isomerase preparation used was the dialyzed ammonium sulfate fraction obtained from sheep kidney cortex as described by Beck *et al.*[3] Its specific activity (0.8–1.0 μmole/hr/mg protein) was a little higher than previously reported. It will be referred to as the enzyme. The apoenzyme was prepared by a modification of the method of Warburg and Christian[15] described by Cori and Illingworth[16] for the removal of bound pyridoxal phosphate from crystalline muscle phosphorylase. In a typical resolution, carried out at 0°, 1.2 ml of enzyme (31.5 mg protein/ml) was diluted to 6.3 mg protein/ml with 1.2 ml of 0.02 *M* Tris buffer, pH 7.4, and 3.6 ml of distilled water. 6.6 ml of saturated (at 0°) ammonium sulfate were then added followed by 1.7 ml of 0.004 *N* HCl. Both additions were made dropwise with vigorous stirring. The resulting pH was 3.45–3.6 (glass electrode). The turbid solution was centrifuged for 7 min at 12,000 rpm in the high speed head of the International refrigerated centrifuge. The precipitate was suspended in 8 ml of 75 per cent saturated ammonium sulfate, previously adjusted to pH 8.7 with NN_4OH, and recentrifuged. This operation was repeated once and the precipitate was then taken up in 1.2 ml of 0.02 *M* Tris buffer, pH 7.4; any insoluble material was removed by centrifugation. This resolution procedure was successfully repeated many times. While the degree of resolution increased with increasing acidity, the yield of undenatured apoenzyme decreased; in our experiments it varied between 35 and 50 per cent.

A purified preparation of methylmalonyl isomerase from *Propionibacterium shermanii* was kindly provided by Dr. H. G. Wood, Western Reserve University School of Medicine. This enzyme was extensively resolved and had but little activity in the absence of added cobamide coenzymes. With saturating concentrations of DBC coenzyme and under our assay conditions,[3] its specific activity was 0.5 μmoles/min/mg protein, 30 to 40 times higher than that of the kidney preparation.

Methylmalonyl isomerase activity was measured in two ways. In most cases (direct assay) the succinyl CoA formed from methylmalonyl CoA as substrate was determined enzymatically as succinate, following alkaline hydrolysis of the thiolester, through reduction of cytochrome *c* in the presence of succinoxidase.[3] In some experiments ($C^{14}O_2$ fixation assay) methylmalonyl isomerase was incubated with propionyl CoA and $C^{14}O_2$, in the presence of an excess of propionyl carboxylase, and the radioactivity incorporated into dicarboxylic acids determined before and after oxidation with $KMnO_4$.[2] The former value gives the amount of methylmalonate plus succinate formed whereas the latter gives that of succinate only as succinate is resistant to permanganate treatment while methylmalonate is destroyed by oxidation to volatile products. Propionyl carboxylase from pig heart[4] (specific activity about 12, free of methylmalonyl isomerase) was provided by Dr. Y. Kaziro of this department. Synthetic propionyl CoA[2] and methylmalonyl CoA[3] were prepared as previously described. Since synthetic methylmalonyl CoA is a racemic mixture of two stereoisomers, only one of which is attacked by the isomerase,[17] the amounts given in the tables are one-half of the total thiolester present. Cytochrome *c*, type III, from horse heart was obtained from the Sigma Chemical Company, St. Louis, Missouri. We are greatly indebted to Dr. Karl Folkers,

Merck Institute for Therapeutic Research, Rahway, New Jersey, for generous supplies of the DBC coenzyme, Dr. H. A. Barker, University of California, Berkeley, California, for gifts of BC and AC coenzymes, and Dr. Leon Ellenbogen, Lederle Laboratories Division, American Cyanamid Company, Pearl River, New York, for gifts of intrinsic factor. The vitamin B_{12}-binding capacity of the highly purified intrinsic factor preparation was 0.5 μg per mg.[18]

Protein was determined spectrophotometrically.[19] Experiments involving the use of cobamide coenzymes were performed in the dark room with a dim red light.

Results.—Effect of DBC coenzyme on methylmalonyl isomerase activity: As shown in Table 1 (experiments 1*a* and 1*b*), DBC coenzyme produced a slight stimulation

TABLE 1

EFFECT OF DIMETHYLBENZIMIDAZOLYLCOBAMIDE COENZYME (DBC) ON ACTIVITY OF SHEEP KIDNEY METHYLMALONYL ISOMERASE (DIRECT ASSAY)

Experiment No.	Enzyme	Additions	Succinate formed (μmoles)
1a	Unresolved	None	0.132
1b	"	DBC (1.5×10^{-5} *M*)	0.214
2a	Resolved	None	0.015
2b	"	DBC (1.6×10^{-6} *M*)	0.127
2c	"	Vitamin B_{12} (1.2×10^{-4} *M*)	0.018
2d	"	DBC + vitamin B_{12}	0.123
2e	Boiled	None	0.007
2f	None	None	0.006
3a	Resolved	None	0.005
3b	"	DBC (4.2×10^{-7} *M*)	0.138
4a	"	None	0.013
4b	"	DBC (10^{-7} *M*)	0.108
4c	"	DBC (10^{-6} *M*) previously illuminated*	0.013
4d	"	DBC (10^{-7} *M*) pretreated with KCN†	0.016
4e	"	DBC (10^{-7} *M*) preincubated with intrinsic factor (2 mg)‡	0.016

Samples contained potassium phosphate buffer, pH 7.4, 50 μmoles; methylmalonyl CoA, 0.35 μmoles and, when present, enzyme with 0.32 to 0.38 mg of protein (except in experiment 3, 0.55 mg). Other additions as indicated. Final volume made up with water to 0.9 ml. Incubation 20 min at 30°. Coenzyme (or vitamin B_{12}) preincubated with enzyme for 12–13 min prior to the addition of methylmalonyl CoA.

* DBC coenzyme (1.8×10^{-5} *M*) illuminated before use in a Beckman cell (d = 1 cm) with a 100-watt tungsten lamp at a distance of 12 cm for 150 min at 0°.

† DBC coenzyme (5×10^{-6} *M*) made 0.1 *N* with KCN and kept for 30 min at room temperature before use.

‡ DBC coenzyme and intrinsic factor were preincubated for 30 min at 30°. Enzyme was then added and the mixture incubated for a further 13 min prior to the addition of methylmalonyl CoA.

of the isomerase activity of unresolved enzyme. This effect, previously noted by Stern and Friedman[20] with ox liver fractions, varied from preparation to preparation and was mostly small. On the other hand, the resolved enzyme had negligible activity in the absence of coenzyme (experiments 2, 3, and 4). Vitamin B_{12} at much higher concentrations was inactive (experiment 2*c*). As expected, prior exposure of the coenzyme to light, cyanide, or intrinsic factor (experiment 4) led to loss of activity. Table 2 shows the effect of the coenzyme on methylmalonyl isomerase activity as determined by the $C^{14}O_2$ fixation assay.

Coenzyme specificity and affinity constants: The rate of methylmalonyl CoA isomerization as a function of the concentration of added DBC or BC coenzyme is shown in Figure 1. Coenzyme concentrations were determined spectrophoto-

metrically at wavelength 367 mμ in 0.1 *N* KCN according to Barker *et al.*[21] Maximal rates with either coenzyme were obtained at concentrations around 3×10^{-6} *M*. On the other hand, 10^{-5} *M* AC coenzyme gave only 15 per cent of the maximal rate, a value given by about 10^{-8} *M* DBC coenzyme. Thus, the effect of the former could have been due to a very slight contamination (0.1 per cent) with the latter.

TABLE 2

EFFECT OF DIMETHYLBENZIMIDAZOLYLCOBAMIDE COENZYME ON ACTIVITY OF SHEEP KIDNEY METHYLMALONYL ISOMERASE ($C^{14}O_2$ FIXATION ASSAY)

Methylmalonyl CoA isomerase	Coenzyme addition	$C^{14}O_2$ fixed (cpm)	
		In methylmalonate	In succinate
Unresolved	None	2820	3400
"	1.5×10^{-5} *M*	1930	3540
Resolved	None	4615*	855*
"	1.5×10^{-5} *M*	2500*	3090*
None	"	5310	90

Samples contained, in μmoles, Tris buffer, pH 7.4, 100; $MgCl_2$, 6; ATP, 3; $KHC^{14}O_3$ (14000 cpm/μmole), 10; propionyl CoA, 0.5; glutathione, 2; propionyl carboxylase, 0.1 unit; unresolved or resolved methylmalonyl isomerase (1 mg protein); and coenzyme (preincubated with isomerase as in Table 1) as indicated. Final volume made up with water to 1.0 ml. Incubation 15 min at 30°.

* Average of duplicate runs.

It is apparent that the AC coenzyme has little if any effect with the kidney methylmalonyl isomerase. A less likely possibility is that this coenzyme is inactivated by the relatively crude enzyme preparation used.

The apparent affinity constants (K_m values), calculated from Lineweaver–Burk[22] plots of the data in Figure 1, are given in Table 3 (column 2). Column 3 gives the similarly obtained values for *P. shermanii* methylmalonyl isomerase.[23] The table

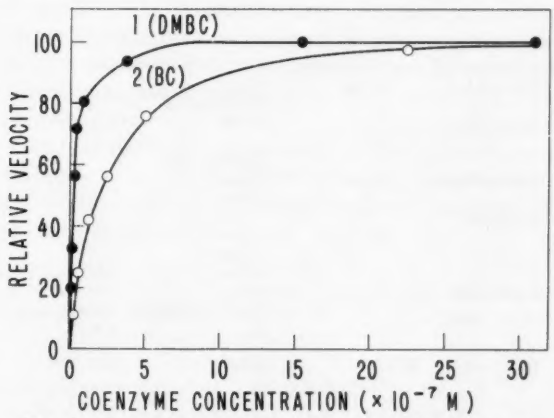


FIG. 1.—Rate of isomerization of methylmalonyl CoA as a function of cobamide coenzyme concentration. Conditions of Table 1 with resolved enzyme. Curve 1, enzyme with 0.9 mg protein; curve 2, enzyme with 0.6 mg protein. Relative velocity given as per cent of V_{max}. DMBC = dimethylbenzimidazolylcobamide coenzyme; BC = benzimidazolylcobamide coenzyme.

also lists for comparison (column 4) the corresponding values of Barker *et al.*[21] for *C. tetanomorphum* glutamic isomerase. It may be seen that both kidney and bacterial methylmalonyl isomerases have highest affinity for the DBC coenzyme while the reverse is true for glutamic isomerase. The affinity for BC coenzyme is about the same in all cases. Curiously enough both bacterial enzymes have high affinity for the AC coenzyme whereas, as noted above, this coenzyme has little or no affinity for the kidney methylmalonyl isomerase.

TABLE 3

COENZYME AFFINITIES OF COBAMIDE ENZYMES

Coenzyme	K_m (M) Kidney methylmalonyl isomerase	K_m (M) *P. shermanii* methylmalonyl isomerase	K_m (M) Glutamic isomerase*
DBC	2.1×10^{-8}	2.4×10^{-8}	1.8×10^{-5}
BC	2.0×10^{-7}	1.3×10^{-7}	2.4×10^{-7}
AC		1.0×10^{-7}	1.4×10^{-6}

Conditions for the experiments with *P. shermanii* methylmalonyl isomerase were those of Table 1 except that only 17 to 20 μg of enzyme protein were used and the incubation was for 15 to 19 min.
* Values from Barker *et al.*[21]

Coenzyme binding strength of native and reconstituted holenzyme: As previously mentioned, such procedures as exposure to light, treatment with charcoal, or incubation with intrinsic factor, which might be expected to bring about resolution of methylmalonyl isomerase, do not significantly affect the activity of the native enzyme (Table 4, experiments 1*a* and *b*, 2*a* and *b*, and 3). However, when applied

TABLE 4

EFFECT OF LIGHT, CHARCOAL TREATMENT, AND INTRINSIC FACTOR ON SHEEP KIDNEY METHYLMALONYL ISOMERASE (DIRECT ASSAY)

Experiment No.	Enzyme	Coenzyme (DBC) addition	Treatment	Succinate formed (μmoles)
1a	Unresolved	None	None	0.126
1b	"	"	Illuminated	0.119
1c	Resolved	"	None	0.032
1d	"	10^{-7} M	"	0.150
1e	"	"	Illuminated	0.070
2a	Unresolved	None	None	0.115
2b	"	"	Charcoal	0.125
2c	Resolved	"	None	0.025
2d	"	10^{-7} M	"	0.167
2e	"	"	Charcoal	0.094
2f	"	"	Charcoal before DBC coenzyme	0.181
3a	Unresolved	None	None	0.110
3b	"	"	Intrinsic factor (1 mg)	0.104
3c	"	"	" " (5 mg)	0.111
3d	"	"	" " (10 mg)	0.098
4a	Resolved	"	None	0.013
4b	"	10^{-7} M	"	0.108
4c	"	"	Intrinsic factor (2 mg)	0.037

Conditions as in Table 1 with the following exceptions. Isomerase (in mg protein) experiment 1, 0.5; experiment 2, unresolved, 0.5, resolved, 0.8; experiment 3, 0.33; experiment 4, 0.39. Preincubation of coenzyme with isomerase, experiment 4, 30 minutes. Illumination, 60 min at 0° in a Beckman cell (d = 1 cm) with two 100 Watt lamps at 12 cm. Charcoal treatment, 1 mg Nuchar C per mg protein added to enzyme solution (2 mg protein/ml), stirred 15 min at 0° and filtered. Protein was redetermined in the supernatant as some is removed by the charcoal. When used, intrinsic factor was incubated with the reaction mixture for 15 min at 30° prior to adding methylmalonyl CoA.

to the reconstituted holoenzyme, i.e., to the resolved enzyme after preincubation with coenzyme, the same procedures lead to partial inactivation (Table 4, experiments 1*c* to *e*, 2*c* to *f*, and 4). This suggests either that the natural coenzyme is not identical with DBC coenzyme, or that holoenzyme binds the coenzyme less tightly in the reconstituted than in the native form. Considering that the enzyme is exposed to a low pH for resolution, the latter is not an unlikely possibility.

Discussion.—The fact that the AC coenzyme has very low, if any activity with kidney methylmalonyl isomerase is not without interest, since this coenzyme is active not only with glutamic isomerase of *C. tetanomorphum*, an organism that contains mainly AC coenzyme, but also with methylmalonyl isomerase of *P. shermanii*, an organism that contains predominantly DBC coenzyme. It may be noted that the behavior of the cobamide coenzymes with the mammalian enzyme parallels that of the corresponding vitamin forms in the animal organism, since both vitamin B_{12} and its benzimidazole analogue are active as growth factors and against pernicious anemia while pseudovitamin B_{12} (the adenine analogue) is inactive.[24]

Eggerer *et al.*[25] have recently established that methylmalonyl CoA isomerization (cf. reaction 1) involves migration of the thiolester rather than the carboxyl[17] group. This is in line with the fact[26] that glutamate isomerization (cf. reaction 2) occurs by migration of a two-carbon fragment, involving carbons 1 and 2 of glutamate, rather than by rearrangement of the C_5-carboxyl group. Rearrangement of a group along a propionic acid residue is thus a common feature of the two reactions; in both cases a methyl group is produced when the straight-chain compound is converted to the branched-chain isomer.

In view of the reported presence of methylmalonic acid in human urine,[27] a study of the levels of this compound in the urine of pernicious anemia patients, and of the effect of vitamin B_{12} thereon, might be of some interest.

Summary.—Methylmalonyl isomerase apoenzyme was prepared from sheep kidney cortex enzyme fractions by acidification in the presence of ammonium sulfate. Activity was restored by dimethylbenzimidazolyl- and benzimidazolylcobamide coenzymes but adenylcobamide coenzyme had little or no effect. In contrast, the three coenzymes were active with methylmalonyl isomerase preparations from *P. shermanii*. A study of the effects of illumination, charcoal treatment and incubation with intrinsic factor on the native and the cobamide coenzyme-reactivated kidney enzyme, suggested that the strength of the apoenzyme-coenzyme binding in this enzyme is greater before than after resolution.

* Paper No. 6 of series "Metabolism of propionic acid in animal tissues." Aided by grant from the National Institute of Arthritis and Metabolic Diseases (Grant A-1845) of the U.S. Public Health Service, and by the Rockefeller Foundation.

[1] Flavin, M., P. J. Ortiz, and S. Ochoa, *Nature*, **176**, 823 (1955).

[2] Flavin, M., and S. Ochoa, *J. Biol. Chem.*, **229**, 965 (1957).

[3] Beck, W. S., M. Flavin, and S. Ochoa, *J. Biol. Chem.*, **229**, 997 (1957).

[4] Tietz, A., and S. Ochoa, *J. Biol. Chem.*, **234**, 1394 (1959).

[5] Abbreviations: ATP, adenosine triphosphate; CoA, coenzyme A; DBC, dimethylbenzimidazolylcobamide; BC, benzimidazolylcobamide; AC, adenylcobamide; Tris, tris(hydroxymethyl)aminomethane.

[6] Swick, R. W., and H. G. Wood, these PROCEEDINGS, **46,** 28 (1960).

[7] Lynen, F., J. Knappe, E. Lorch, G. Jütting, and E. Ringelmann, *Angew. Chem.*, **71,** 481 (1959).

[8] Kaziro, Y., E. Leone, and S. Ochoa, 137th meeting, American Chemical Society, Cleveland, Ohio, April, 1960; Abstracts of papers, p. 39C; these PROCEEDINGS, **46,** 1319 (1960).

[9] Barker, H. A., R. D. Smyth, E. J. Wawszkiewicz, M. N. Lee, and R. M. Wilson, *Arch. Biochem. Biophys.*, **78,** 468 (1958).

[10] For simplicity the name methylmalonyl isomerase will be used for the enzyme (or enzymes) catalyzing reaction 1 and glutamic isomerase for the enzyme (or enzymes) catalyzing reaction 2, without implication that a single enzyme is involved in each case, a matter which is as yet unsettled. The former name, as are the names β-methylcrotonyl carboxylase and propionyl carboxylase, is in line with current nomenclature of enzymes of fatty acid metabolism (Beinert, H., D. E. Green, P. Hele, O. Hoffmann-Ostenhof, F. Lynen, S. Ochoa, G. Popják, and R. Ruyssen, *Science,* **124,** 617 (1956)).

[11] Barker, H. A., H. Weissbach, and R. D. Smyth, these PROCEEDINGS, **44,** 1093 (1958); Weissbach, H., J. J. Toohey, and H. A. Barker, these PROCEEDINGS, **45,** 521 (1959).

[12] Smith, R. M., and K. J. Monty, *Biochem. Biophys. Res. Comm.*, **1,** 105 (1959).

[13] Gurnani, S., S. P. Mistry, and B. C. Johnson, *Biochim. et Biophys. Acta*, **38,** 187 (1960).

[14] Stadtman, E. R., P. Overath, H. Eggerer, and F. Lynen, *Biochem. Biophys. Res. Comm.*, **2,** 1 (1960).

[15] Warburg, O., and W. Christian, *Biochem. Z.*, **298,** 150 (1938).

[16] Cori, C. F., and B. Illingworth, these PROCEEDINGS, **43,** 547 (1957).

[17] Beck, W. S., and S. Ochoa, *J. Biol. Chem.*, **232,** 931 (1958).

[18] Ellenbogen, L., and W. L. Williams, *Biochem. Biophys. Res. Comm.*, **2,** 340 (1960).

[19] Warburg, O., and W. Christian, *Biochem. Z.*, **310,** 384 (1942).

[20] Stern, J. R., and D. L. Friedman, *Biochem. Biophys. Res. Comm.*, **2,** 82 (1960).

[21] Barker, H. A., R. D. Smyth, H. Weissbach, J. I. Toohey, J. N. Ladd, and B. E. Volcani, *J. Biol. Chem.*, **235,** 480 (1960).

[22] Lineweaver, H., and D. Burk, *J. Am. Chem. Soc.*, **56,** 658 (1934).

[23] These values are in good agreement with values obtained by E. R. Stadtman (personal communication to Dr. H. A. Barker) with a different preparation of *P. shermanii* methylmalonyl isomerase.

[24] Coates, M. E., and J. E. Ford, in *The Biochemistry of Vitamin* B_{12}, Biochemical Society Symposia No. 13 (Cambridge University Press, 1955), p. 36.

[25] Eggerer, H., P. Overath, F. Lynen, and E. R. Stadtman, *J. Am. Chem. Soc.*, **82,** 2643 (1960). Similar results have been obtained by Phares, E. F., M. V. Long, and S. F. Carson with methylmalonyl isomerase of *Propionibacterium pentosaceum.* 138th meeting, American Chemical Society, New York, September 1960, Abstracts of papers, p. 21C.

[26] Munch-Petersen, A., and H. A. Barker, *J. Biol. Chem.*, **230,** 649 (1958).

[27] Thomas, K., and K. Stalder, *Chem. Ber.*, **90,** 970 (1957).

*BIOTIN AND PROPIONYL CARBOXYLASE**

By Yoshito Kaziro,† Enzo Leone, ‡and Severo Ochoa

DEPARTMENT OF BIOCHEMISTRY, NEW YORK UNIVERSITY SCHOOL OF MEDICINE

Communicated August 10, 1960

Propionyl carboxylase catalyzes the Mg^{++}-dependent reversible carboxylation of propionyl CoA[1] to methylmalonyl CoA according to reaction 1.[2] In view

$$\text{ATP} + \text{CO}_2 + \begin{matrix}\text{CH}_3\\|\\\text{CH}_2\\|\\\text{COSCoA}\end{matrix} \overset{(Mg^{++})}{\rightleftharpoons} \text{ADP} + \text{P}_i + \begin{matrix}\text{CH}_3\\|\\\text{HC—COOH}\\|\\\text{COSCoA}\end{matrix} \qquad (1)$$

of the widespread indications that biotin functions in carboxylation reactions and, more specifically, of the observation by Lardy and Adler[3] that mitochondrial extracts of biotin-deficient rat liver carboxylated propionate at a greatly reduced rate, highly purified preparations of propionyl carboxylase from pig heart were assayed for biotin;[2] however, none was detected at that time. A reinvestigation of this problem was prompted by the finding that preparations of acetyl[4] and β-methylcrotonyl[5] carboxylase from liver and a bacterial source, respectively, contained biotin and were inactivated by avidin, the biotin-binding protein of egg white. The reactions catalyzed by these enzymes—carboxylation of acetyl CoA to malonyl CoA and of β-methylcrotonyl CoA to β-methylglutaconyl CoA—conform to the pattern of reaction 1.[6] It appeared, therefore, that propionyl carboxylase must contain biotin. Reinvestigation of this question was the more desirable as our biotin determinations had been carried out by bioassay with *Lactobacillus arabinosus* while the more sensitive, but less specific yeast bioassay had been used for analysis of other carboxylation enzymes.[5]

Highly purified pig heart propionyl carboxylase has now been found to contain 1 mole of bound biotin per 1.5 to 1.6 × 10^5 gm of protein. Since the molecular weight of this enzyme[7] is 1.5 × 10^6, one molecule contains 10 molecules of biotin. Activity of the enzyme is inhibited by stoichiometric amounts of avidin. Previous indications[2] that the enzyme depends on the integrity of -SH groups for activity have been confirmed and extended by the observation that it is highly sensitive to sulfhydryl-binding reagents.

Final proof for the involvement of a single enzyme has been provided by the crystallization of pig heart propionyl carboxylase.[8] The preparation and properties of the crystalline enzyme, as well as experiments on the carboxylation and decarboxylation of the enzyme itself, will be reported later. However, observations on the propionyl carboxylase-catalyzed exchange of radioactive orthophosphate and ADP with ATP, bearing on the reaction mechanism, will be presented here. In confirmation of previous results[2] they rule out the reported occurrence of separate steps involving reaction of ATP with the enzyme[5,9] to form either ADP and an enzyme-phosphate or phosphate and an enzyme-ADP compound in the absence of CO_2.

Materials and Methods.—Propionyl carboxylase was assayed as described by Tietz and Ochoa[2] either by measuring the fixation of $C^{14}O_2$ or by the optical assay, except

that the latter was conducted at 25° rather than at 30°. One unit was defined as the amount of enzyme catalyzing the fixation of 1 μmole of CO_2, or the production of 1 μmole of ADP per min at 25°. The enzyme was prepared from pig heart as described by Tietz and Ochoa[2] with an additional step involving chromatography on a hydroxylapatite column.[10] The best amorphous preparations had a specific activity of 12.5 units/mg protein. At 30° the specific activity is 1.5 times higher or 18.75. Ultracentrifugal analysis[7] revealed the presence of two main components (Fig. 1). The enzyme, the fast moving component ($s_{20,w}$ = 18.02 S), represented 75–80 per cent of the total protein. Biotin was determined by microbiological assay after acid hydrolysis using both the yeast[11] and the *L. arabinosus*[12] method. Cultures of *Saccaromyces cerevisiae* (ATCC No. 9896) and *L. arabinosus* (ATCC *L. plantarum*, No. 8014) were obtained from the American Type Culture Collection, Washington, D. C.

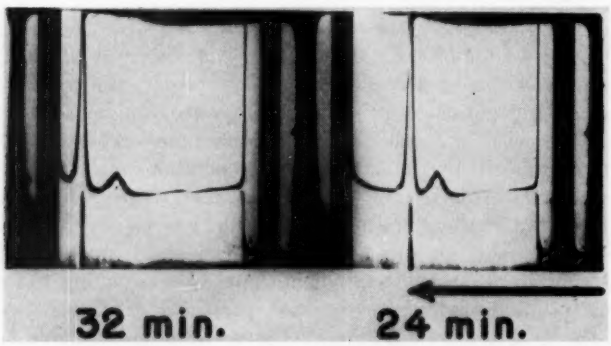


FIG. 1.—Sedimentation patterns of propionyl carboxylase, specific activity 12.5. Sedimentation at 59, 780 rpm in 0.05 *M* tris buffer, pH 7.5, containing 0.001 *M* GSH. Protein concentration, 0.45%. Main component, s_{20}, W = 18.02 S.

To measure the exchange of P^{32}-labeled orthophosphate or ADP with ATP, the reaction was stopped by addition of an equal volume of 10 per cent perchloric acid followed by adsorption of the nucleotides on Norit and elution with ethanol-ammonia.[13] In the experiments on P_i^{32} exchange of Table 2, aliquots of the eluate were counted directly. In all other cases elution was followed by paper chromatographic separation of the nucleotides in the isobutyric acid-ammonia system[14] with use of Whatman No. 1 paper. The ATP spots were cut out and eluted with 0.1 *N* HCl, and the radioactivity of the eluates was determined with an end-window counter. In the experiments on simultaneous exchange of P_i^{32} and ADP-C^{14} with ATP, the reaction was stopped with 0.2 ml of 20 per cent trichloroacetic acid and the nucleotides were isolated as above by paper chromatography of the Norit eluates. To determine P^{32} and C^{14} radioactivity separately, the isolated ATP was hydrolyzed in 1.0 *N* HCl for 10 min at 100°. This liberates the labile phosphate groups—of which the terminal is labeled with P^{32}—as orthophosphate and the C^{14}-labeled nucleotide moiety partly as AMP, partly as adenine and ribose 5-phosphate. The hydrolysate was passed through a Norit column, previously washed with secondary octanol,[15] to adsorb adenine and AMP. The effluent containing the orthophosphate

was collected and the nucleotide fraction was eluted with ethanol-ammonia. Adenine and AMP were determined spectrophotometrically at wavelength 257 mμ; phosphate was determined colorimetrically.[16] C^{14}-radioactivity was measured with a gas-flow, windowless counter.

P^{32}-labeled ADP was prepared from ATP equally labeled in the middle and terminal phosphate groups by reaction with glucose in the presence of hexokinase.[17] Crystalline yeast hexokinase was a gift of Dr. M. Kunitz, The Rockefeller Institute, New York. The labeled ATP was prepared as described by Tatibana *et al.*[18] In this method labeled ATP is formed by glycolyzing human blood cells suspended in a medium containing citrate, glucose and P_i^{32}. The ATP^{32} was purified by ion-exchange chromatography[19] on a Dowex 1, X-2 (Cl^- form, 200–400 mesh) column and concentrated by chromatography on a small column of the same resin. It was then precipitated as the barium salt, converted to the free acid by shaking with Dowex 50, X-4 (H^+ form, 50–100 mesh), neutralized to pH 3.6 with NaOH and lyophilized. The specific radioactivity was 2×10^7 cpm/μmole. The ADP^{32} obtained from the hexokinase reaction was purified in the same way as the ATP, except that it was diluted 1:10 with non-labeled carrier ADP before chromatography. The final specific radioactivity of the ADP was 10^6 cpm/μmole.

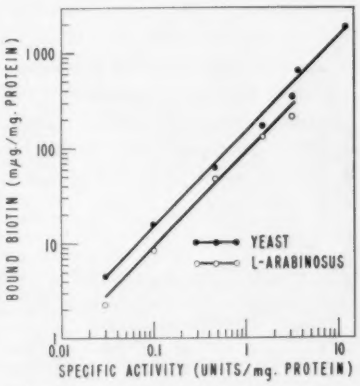


FIG. 2.—Proportionality between biotin content and specific activity of pig heart propionyl carboxylase. Biotin assayed by the yeast (—●—●—●—) and *L. arabinosus* (—○—○—○—) methods.

Propionyl CoA was prepared as in previous work.[2] Avidin (2.5 units/mg) was purchased from Nutritional Biochemicals Corp., Cleveland, Ohio, and crystalline pyruvic kinase from C. F. Boehringer and Sons, Mannheim, Germany. We are indebted to Dr. W. E. Scott, Hoffmann-LaRoche, Inc., Nutley, New Jersey, for a gift of *d*-biotin and to Dr. Karl Folkers, Merck Institute for Therapeutic Research, Rahway, New Jersey, for a gift of biocytin. The source of all other preparations, including 8-C^{14}-ADP, was as previously given.[2]

Results.—Biotin content and avidin inhibition: Propionyl carboxylase was assayed for biotin after acid hydrolysis at all steps of purification. As shown in Figure 2, the biotin content increased in proportion to the degree of purification. As determined by the yeast assay, preparations of specific activity 12.5 contained 1.6 μg of bound biotin/mg protein, corresponding to one mole/156,000 gm protein. The *L. arabinosus* assay gave on the average 30 per cent lower values. This assay was more sensitive to variations in pH and salt concentration than the yeast assay, a factor which may have been responsible for the previous failure to detect the presence of biotin in this enzyme.[2] No biotin was detected by either assay prior to hydrolysis. Such a result could be obtained either if biotin were dissociably bound to the enzyme with a very low dissociation constant, or if it were part of a larger coenzyme molecule even if this were loosely bound to the protein. However, both

these possibilities were ruled out by the observation that the biotin content of the enzyme remained unchanged after precipitation with perchloric acid, a procedure which would have split off any compound not covalently bound to the protein. Since biotin is bound through a peptide-like linkage to the epsilon-amino group of lysine in biocytin,[20] it is likely that the vitamin is thus linked to lysine residues of the enzyme.

As reported for other biotin enzymes[4,5] and for propionyl carboxylase of ox liver,[9] pig heart propionyl carboxylase was inhibited by avidin. As would be expected, the inhibition could be prevented by the prior addition of excess biotin. However, as also reported by Halenz and Lane,[9] biotin addition did not restore the activity of enzyme previously inhibited by avidin. The activity of avidin solutions decreased with time, even on storage in the frozen state. With freshly prepared solutions, however, close to 100 per cent inhibition was obtained with stoichiometric amounts of avidin after incubation with the enzyme for 10–15 min (Fig. 3). This indicates

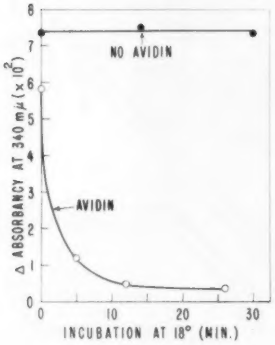


FIG. 3.—Time course of inhibition of propionyl carboxylase by a stoichiometric amount of avidin. 0.12 ml. of a freshly prepared solution of avidin (2.5 units/mg) with 0.03 mg of protein (0.075 unit*) was incubated at 18° with 40 μg of enzyme† (0.075 μg bound biotin) in 0.02 M Tris buffer, pH 7.5, containing 0.001 *M* EDTA and 0.0005 *M* GSH. Final volume, 0.6 ml. At zero time and at various time intervals thereafter, 0.04 ml aliquots (equivalent to 2.66 μg enzyme) were withdrawn for propionyl carboxylase activity determination by the optical assay. In the similarly treated control sample (no avidin) buffer was substituted for the avidin solution. The ordinate gives reaction rates as Δ absorbancy/min.

* 1 unit is the amount of avidin binding 1 μg of free biotin.[21] The protein concentration of the avidin solution was determined by the method of Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall, *J. Biol. Chem.*, **193**, 265 (1951).

† Its specific activity, initially 12.5, had decreased to 7.5 at the time the enzyme was used for this experiment.

that avidin has as much affinity for enzyme-bound biotin as it has for the free vitamin a result which is in line with the observation of Wright *et al.*[22] that avidin has the same degree of affinity for biotin and biocytin. The purest preparations of avidin have a binding capacity of 8 μg biotin/mg (8 units[21]/mg) and a molecular weight between 60,000 and 70,000.[23] Since their biotin-binding capacity corresponds to one mole biotin/30,000 gm avidin, Warner[24] considers it likely that two molecules of biotin are bound by one of avidin. If so, for a molecular weight of 1.5×10^6, one molecule of pig heart propionyl carboxylase containing 10 molecules of biotin would bind 5 molecules of avidin.

With bacterial β-methylcrotonyl carboxylase Lynen *et al.*[5] obtained carboxylation of *d*-biotin when added in large amounts to the system in place of β-methylcrotonyl CoA. *l*-Biotin, the non-natural enantiomorph, did not react. Under similar conditions, and with large amounts of pig heart propionyl carboxylase (up to 750 μg protein/ml; specific activity, 4), we observed no detectable reaction either in the optical assay or in experiments with $C^{14}O_2$. Biocytin also gave negative results. The results of Lynen *et al.* could be explained by the assumption[5] that their enzyme can bind *d*-biotin in a dissociable fashion and that the CO_2 combined with (covalently bound) enzyme biotin can be transferred to the dissociably bound vitamin. If so,

our negative results would simply reflect a lack of affinity of propionyl carboxylase for free biotin. Another possibility would be that the relatively crude enzyme preparations used by Lynen *et al.* in these experiments contained enzymes capable of hydrolyzing biotin (or biotin-CO_2) off the carboxylase and of linking biotin to the enzyme in the presence of ATP.

Inhibition by -SH reagents: It was noticed in previous work[2] that GSH not only stabilized pig heart propionyl carboxylase but also enhanced the rate of reaction suggesting the involvement of -SH groups. For this reason, the effect of such sulfhydryl-binding reagents as *p*-chloromercuribenzoate and iodoacetamide was investigated with use of the optical assay. As shown in Table 1, *p*-chloromercuribenzoate was strongly inhibitory; complete inhibition was produced by a concentration of 5×10^{-6} *M*. Under our experimental conditions (brief preincubation with the enzyme) iodoacetamide was less inhibitory; 90 per cent inhibition was obtained at a concentration of 10^{-2} *M*. The same results were obtained with *p*-chloromercuribenzoate in the $C^{14}O_2$ fixation assay thus ruling out an effect of the inhibitor on the auxiliary enzymes used in the optical assay. 0.5×10^{-3} *M* propionyl CoA, added prior to the sulfhydryl reagent, afforded partial protection (about 60 per cent) of the inhibition produced by 5×10^{-6} *M* *p*-chloromercuribenzoate or 10^{-2} *M* iodoacetamide. The same was true of 2×10^{-3} *M* GSH. Under the same conditions, 2×10^{-3} *M* ATP provided no significant protection. Inhibition with *p*-chloromercuribenzoate could be partially reversed (about 50 per cent) by 4×10^{-3} *M* GSH.

TABLE 1

INHIBITION OF PROPIONYL CARBOXYLASE BY *p*-CHLOROMERCURIBENZOATE AND IODOACETAMIDE

p-Cl-Hg-benzoate (*M*)	Inhibition (%)	Iodoacetamide (*M*)	Inhibition (%)
1×10^{-7}	31	1×10^{-4}	34
5×10^{-7}	57	1×10^{-3}	50
1×10^{-6}	77	1×10^{-2}	89
2×10^{-6}	94		
5×10^{-6}	100		

Following incubation of the enzyme with the inhibitor for 5 min at 18°, the reaction was started by addition of ATP and propionyl CoA. Optical assay. Samples contained (in μmoles): Tris buffer, pH 8.0, 100; $MgCl_2$, 4; ATP, 2; $KHCO_3$, 50; propionyl CoA, 0.5; KCl, 100; phosphoenolpyruvate, 1; DPNH, 0.16; an excess of pyruvic kinase and lactic dehydrogenase; and propionyl carboxylase (specific activity, 12.5) 0.8 μg.

Experiments on reaction mechanism: On the basis of experiments on the exchange of radioactive P_i[5] or ADP[9] with ATP, the exchange of β-methylcrotonyl CoA with β-methylglutaconyl CoA[5] and the carboxylation of free biotin,[5] it has been proposed[5, 9] that the enzymatic carboxylation of β-methylcrotonyl CoA or propionyl CoA involves three separate steps as follows:[25]

(1*a*) ATP + enzyme-biotin $\overset{(Mg^{++})}{\rightleftharpoons}$ P_i + enzyme-biotin∼ADP (or ADP + enzyme-biotin∼P_i)

(1*b*) Enzyme-biotin∼ADP (or enzyme-biotin∼P_i) + CO_2 $\overset{(Mg^{++})}{\rightleftharpoons}$ ADP (or P_i) + enzyme-biotin∼CO_2

(1*c*) Enzyme-biotin∼CO_2 + β-methylcrotonyl CoA (or propionyl CoA) $\rightleftharpoons$ enzyme-biotin + β-methylglutaconyl CoA (or methylmalonyl CoA)

Failure to obtain exchange of either P_i^{32} or ADP-C^{14} with ATP in previous work with pig heart propionyl carboxylase[2] made the occurrence of reaction 1*a*, in either of its two forms, very doubtful. In reinvestigating this point we found that there

was hardly any exchange of P_i^{32} with ATP unless ADP was also present (Table 2) and, conversely, no significant exchange of radioactive ADP with ATP in the absence of P_i (Table 3). In all cases the exchange required Mg^{++} and was inhibited by avidin and by *p*-chloromercuribenzoate. However, complete inhibition with the latter was obtained only at 10^{-4} *M* concentration. These results definitely rule out reaction 1*a*. It should be pointed out that when first repeating the experiments on exchange of ADP-C^{14} with ATP we had positive results in the absence of added P_i but this could be traced to the presence of P_i in the ADP and ATP preparations used. As seen in Table 3 there was no significant exchange of ADP-C^{14} with ATP, in the absence of added P_i, when the nucleotides were purified by ion-exchange chromatography prior to use. Since both ATP and ADP preparations slowly decompose on storage (even in the cold and in solid form) to ADP + P_i and AMP + P_i, contamination of ATP with ADP and P_i and of ADP with P_i must have been responsible for the diverging results of other investigators.

TABLE 2

P_i^{32}-ATP EXCHANGE

Experiment 1 (pH 7.5)		Experiment 2 (pH 6.2)	
Components	Radioactivity in nucleotides (cpm)	Gas phase	Radioactivity in nucleotides (cpm)
Complete system	14300	Nitrogen	90
No enzyme	45	CO_2	604
No ADP	1450		
No Mg^{++}	130		
Complete + avidin (0.45 mg)	60		

In experiment 1 the complete system consisted of the following (in μmoles). Tris buffer, pH 7.5, 100; $MgSO_4$, 6; GSH, 2; ATP, 2; ADP, 2; $KH_2P^{32}O_4$, 2 (350,000 cpm); enzyme, 0.54 unit. Final volume, 1.0 ml. In experiment 2 each sample contained (in μmoles) cacodylate buffer, pH 6.2, 200; $MgSO_4$, 6; GSH, 2; ATP, 2; ADP, 2; $KH_2P^{32}O_4$, 2 (516,000 cpm); enzyme, 0.4 unit. Final volume, 1.0 ml. Before adding the enzyme the reaction mixtures were acidified to pH 3.0 with 2.0 *N* HCl, gassed with a stream of nitrogen for 10 min and then brought to pH 6.2 with CO_2-free NaOH. The reaction was carried out in Warburg vessels one of which contained 0.1 ml of 20 per cent KOH in the center well and nitrogen in the gas phase, while the other had CO_2 in the gas phase. Both vessels were shaken for 15 min at 37° before starting the reaction by tipping in the enzyme from the side-bulb. In both experiments 1 and 2 the incubation was for 60 min at 37°.

TABLE 3

ADP^{32}-ATP EXCHANGE

Components	Radioactivity in ATP (cpm/μmole)
Complete system*	74900
No P_i	1350
No Mg^{++}	0
Complete system + avidin (0.45 mg)	3450

* Contained (in μmoles) Tris buffer, pH 7.5, 50; $MgSO_4$, 3; GSH, 1; ATP, 1; ADP (labeled with P^{32} in the terminal phosphate, specific radioactivity, 571,000 cpm/μmole), 0.9; potassium phosphate, pH 7.5, 1; enzyme, 0.25 unit. Final volume, 0.5 ml. Incubation, 60 min at 37°. ATP and ADP^{32} were purified by ion-exchange chromatography prior to their use for this experiment.

Reaction 1*a* is further ruled out by the demonstration that CO_2 is required for the ADP-dependent P_i^{32}-ATP exchange. This is shown in Table 2 (experiment 2). In this experiment efforts were made to eliminate CO_2 as far as possible from the reaction mixture. To further insure absence of CO_2, the reaction was carried out at pH 6.2. Although at this pH propionyl carboxylase has only about 10 per cent of its activity at pH 7.5, the marked stimulation of the exchange by CO_2 is clearly seen. Some stimulation of the simultaneous exchange of P_i^{32} and ADP-C^{14} with ATP was found on increasing the concentration of bicarbonate from 5×10^{-3} *M* (K_m for CO_2, 2.5×10^{-3} *M*[2]) to 10^{-2} *M* (Table 4). Higher concentrations were inhibitory.

The experiments of Table 4 indicate that, as followed simultaneously, each exchange, i.e. that of P_i with ATP and of ADP with ATP, proceeds at about the same rate and each is similarly affected by changes in the concentration of bicarbonate. These results provide additional support for a concerted reaction mechanism since independent steps, as in reactions 1*a* and 1*b*, would probably give different rates of exchange for P_i and ADP.

TABLE 4

SIMULTANEOUS EXCHANGE OF P^{32}-LABELED PHOSPHATE AND C^{14}-LABELED ADP WITH ATP

$KHCO_3$ addition (M)	P_i^{32}-ATP exchange cpm/μmole ATP	%	ADP-C^{14}-ATP exchange cpm/μmole ATP	%
5×10^{-3}	42610	9	14680	6
1×10^{-2}	50180	11	21580	9
2×10^{-2}	47380	10	16680	7
5×10^{-2}	22980	5	10480	4

Samples contained (in μmoles) Tris buffer, pH 7.5, 50; $MgSO_4$, 3; GSH, 1; ATP, 1; 8-C^{14}-ADP (240,000 cpm/μmole) 1; $KH_2P^{32}O_4$ (480,000 cpm/μmole), 1; $KHCO_3$ as indicated; enzyme (specific activity, 12.5), 0.3 unit. Final volume, 0.5 ml. Incubation, 60 min at 37°.

In further confirmation of previous results[26] we failed to obtain an exchange between C^{14}-labeled propionyl CoA and methylmalonyl CoA in the presence of propionyl carboxylase, whether in the absence or presence of Mg^{++}. This rules out reaction 1*c*.

Discussion.—The above results and experiments now under way on the carboxylation and decarboxylation of the propionyl carboxylase protein[8] throw some new light on the mechanism of reaction 1. It is suggested that it proceeds via the separate steps represented by reactions A and B (Fig. 4) and that a side-reaction (reaction A1) is responsible for the observed CO_2-dependent exchange of P_i^{32} and ADP^{32} (or ADP-C^{14}) with ATP, of which the former is ADP- and the latter P_i-dependent.

$$\begin{array}{l} \text{ATP + enzyme-biotin-}CO_2 \\ \qquad\qquad \Big\uparrow\Big\downarrow \text{(A2)} \quad (Mg^{++}) \\ \text{(A) ATP} + CO_2 + \text{enzyme-biotin} \underset{}{\overset{(Mg^{++})}{\rightleftharpoons}} \text{ATP-enzyme-biotin-}CO_2 \overset{(Mg^{++})}{\rightleftharpoons} \left[\begin{matrix}\text{ADP} \\ P_i\end{matrix}\right] \text{enzyme-biotin}\sim CO_2 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \Big\uparrow\Big\downarrow \text{(A1)} \quad (Mg^{++}) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{ADP + } P_i \text{ + enzyme-biotin} \sim CO_2 \\ \text{(B)} \left[\begin{matrix}\text{ADP} \\ P_i\end{matrix}\right] \text{enzyme-biotin} \sim CO_2 + \text{propionyl CoA} \overset{(Mg^{++})}{\rightleftharpoons} \text{ADP + } P_i \text{ + enzyme-biotin +} \\ \text{methylmaloyl CoA} \end{array}$$

FIG. 4.—Suggested mechanism of propionyl carboxylase reaction. The sign ~ denotes a high energy bond.

Although this point requires further documentation, these exchanges appear to be too slow to account for the rate of reaction 1 even in the back direction[2] in which the reaction is much slower than in the forward direction.

Reactions A and B (Fig. 4) have been separated in time by (*a*) stoichiometric carboxylation of the protein with $C^{14}O_2$ (one mole per mole of biotin), in the presence of ATP and Mg^{++}, and (*b*) transfer of the $C^{14}O_2$ from the labile carboxyl-enzyme to

propionyl CoA. When the C^{14}-carboxyl-enzyme is obtained free of nucleotides by passage through a Dowex-1 column, its radioactivity is quantitatively transferred to propionyl CoA, in the presence of a large excess of unlabeled bicarbonate, only after addition of ATP, but there is no transfer if EDTA is added in excess, showing that Mg^{++} is required along with ATP for this step. Formation of "active" carboxyl-enzyme (represented as $\begin{bmatrix} ADP \\ P_i \end{bmatrix}$ enzyme-biotin$\sim Co_2$ in Fig. 4) under these conditions might occur via reaction A2. These results explain our failure to obtain exchange of C^{14}-labeled propionyl CoA with methylmalonyl CoA on incubation with propionyl carboxylase, an exchange that should take place if reaction 1*c* (p. 1323) were indeed a partial step of reaction 1 (p. 1319). Experiments with O^{18} now in progress[27] and determination of the stability of the C^{14}-carboxyl-enzyme, in the absence and presence of ATP, may lead to a closer understanding of the reaction mechanism and the chemical nature of the active form of the CO_2-enzyme.

Repeated observations that certain CO_2-fixation reactions or enzymes were decreased in biotin deficiency had long pointed to a function of this vitamin in CO_2 fixation,[28] but its mode of action was not disclosed until recently by Lynen's decisive work on β-methylcrotonyl carboxylase.[5] Noteworthy is the fact that the "malic" enzyme, the first CO_2 fixation enzyme that was found to be markedly and specifically decreased in biotin deficiency,[29] is not a biotin enzyme. Partially purified preparations of this enzyme had essentially no biotin[29] and, as we observed more recently, the enzyme was not inhibited by avidin. Whether or not biotin is in some way connected with the synthesis of "malic" enzyme, as suggested by Blanchard *et al.*,[30] the relationship of biotin to this enzyme is an intriguing subject for further investigation.

Summary.—Highly purified pig heart propionyl carboxylase contains approximately one mole of bound biotin per 150,000 gm of enzyme. No biotin is released from the protein following precipitation with perchloric acid. This indicates that the vitamin is covalently bound to the protein, probably through peptide-like linkages to the epsilon-amino groups of lysine residues as in biocytin. After short incubation, propionyl carboxylase is completely inhibited by stoichiometric amounts of avidin, the biotin-binding protein of egg white, indicating that this protein has as much affinity for enzyme-bound biotin as it does for the free vitamin or for biocytin biotin. The enzyme is strongly inhibited by *p*-chloromercuribenzoate and less strongly by iodoacetamide indicating that the integrity of -SH groups is essential for activity.

Experiments on the propionyl carboxylase-catalyzed, avidin- and *p*-chloromercuribenzoate-sensitive exchange of radioactive phosphate and ADP with ATP rule out the reported occurrence of steps involving reaction of ATP with the enzyme, to form either ADP and an enzyme-phosphate or phosphate and an enzyme-ADP intermediate in the absence of CO_2, as part of the over-all reaction. The reaction mechanism is discussed in the light of these and other experiments now under way on the carboxylation and decarboxylation of the enzyme itself.

We wish to thank Mr. Morton C. Schneider who carried out most of the biotin assays. We are also indebted to Mr. Schneider and Mr. Horace Lozina for help with the preparation of the enzyme.

* Paper No. 7 of series "Metabolism of propionic acid in animal tissues." Preliminary communication, Symposium on Metabolic Reactions Involving Biotin, 137th meeting, American Chemical Society, Cleveland, Ohio, April 1960; Abstracts of papers, p. 39C. Aided by grants from the National Institute of Arthritis and Metabolic Diseases (Grant A-1845) of the U.S. Public Health Service, and the Rockefeller Foundation.

† Foreign Postdoctoral Fellow of the National Institutes of Health, U.S. Public Health Service.

‡ Rockefeller Foundation Fellow on leave of absence from the University of Sassari, Italy. Permanent address: Zoological Station, Naples, Italy.

[1] Abbreviations: CoA, coenzyme A; ATP, ADP and AMP, adenosine 5′-triphosphate, diphosphate and monophosphate; P_i, inorganic phosphate; DPNH, reduced diphosphopyridine nucleotide; GSH, reduced glutathione; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediamine tetraacetate.

[2] Tietz, A., and S. Ochoa, *J. Biol. Chem.*, **234,** 1394 (1959).

[3] Lardy, H. A., and J. Adler, *J. Biol. Chem.*, **219,** 933 (1956).

[4] Wakil, S. J., E. B. Titchener, and D. M. Gibson, *Biochim. et Biophys. Acta*, **29,** 225 (1958); Wakil, S. J., and D. M. Gibson, *Biochim. et Biophys. Acta*, **41,** 122 (1960).

[5] Lynen, F., J. Knappe, E. Lorch, G. Jütting, and E. Ringelmann, *Angew. Chem.*, **71,** 481 (1959).

[6] The terminology used here for the enzymes catalyzing the carboxylation of acetyl, propionyl, and β-methylcrotonyl CoA is in accordance with currently accepted nomenclature of enzymes of fatty acid metabolism (Beinert, H., D. E. Green, P. Hele, O. Hoffmann-Ostenhof, F. Lynen, S. Ochoa, G. Popják, and R. Ruyssen, *Science*, **124,** 617 (1956).

[7] Warner, R. C., and J. Y. Chen, unpublished results.

[8] Kaziro, Y., and S. Ochoa, unpublished.

[9] Halenz, D. R., and M. D. Lane, *J. Biol. Chem.*, **235,** 878 (1960).

[10] Tiselius, A., S. Hjerten, and O. Levin, *Arch. Biochem. Biophys.*, **65,** 132 (1956).

[11] Lichstein, H. C., *J. Biol. Chem.*, **212,** 217 (1955).

[12] Wright, L. D., and H. R. Skeggs, *Proc. Soc. Exp. Biol. Med.*, **56,** 95 (1944).

[13] Crane, R. K., and F. Lipmann, *J. Biol. Chem.*, **201,** 235 (1953).

[14] Krebs, H. A., and R. Hems, *Biochim. et Biophys. Acta*, **12,** 172 (1953).

[15] Threlfall, C. J., *Biochem. J.*, **65,** 694 (1957).

[16] Lohmann, K., and L. Jendrassik, *Biochem. Z.*, **178,** 419 (1926).

[17] Heppel, L. A., P. J. Ortiz, and S. Ochoa, *J. Biol. Chem.*, **229,** 695 (1957).

[18] Tatibana, M., M. Nakao, K. Miyamoto, T. Sekiguchi, and H. Yoshikawa, *J. Biochem. (Tokyo)*, **46,** 711 (1959).

[19] Cohn, W. E., and C. E. Carter, *J. Am. Chem. Soc.*, **72,** 4273 (1950).

[20] Wright, L. D., E. L. Cresson, H. R. Skeggs, T. R. Wood, R. L. Peck, D. E. Wolf, and K. Folkers, *J. Am. Chem. Soc.*, **74,** 1996 (1952).

[21] Eakin, R. E., E. E. Snell, and R. J. Williams, *J. Biol. Chem.*, **140,** 535 (1941).

[22] Wright, L. D., E. L. Cresson, K. V. Liebert, and H. R. Skeggs, *J. Am. Chem. Soc.*, **74,** 2004 (1952).

[23] Fraenkel-Conrat, H., N. S. Snell, and E. D. Ducay, *Arch. Biochem. Biophys.*, **39,** 80 (1952).

[24] Warner, R. C., in *The Proteins*, ed. H. Neurath and K. Bailey (New York: Academic Press, 1953), vol. IIA, p. 435.

[25] Enzyme-biotin denotes enzyme with covalently bound biotin; the sign ∼ in biotin∼ADP, biotin∼P_i and biotin∼CO_2, denotes a high-energy bond.

[26] Tietz, A., and S. Ochoa, unpublished results.

[27] In collaboration with Dr. Paul D. Boyer, University of Minnesota, Minneapolis, Minnesota.

[28] Lardy, H. A., in Symposium on Metabolic Reactions Involving Biotin, 137th meeting, American Chemical Society, Cleveland, Ohio, April 1960; Abstracts of papers, p. 37C.

[29] Ochoa, S., A. Mehler, M. L. Blanchard, T. H. Jukes, C. E. Hoffmann, and M. Regan, *J. Biol. Chem.*, **170,** 413 (1947).

[30] Blanchard, M. L., S. Korkes, A. del Campillo, and S. Ochoa, *J. Biol. Chem.*, **187,** 875 (1950).

*THE FUNCTION OF ACETATE IN PHOTOSYNTHESIS BY GREEN BACTERIA**

By W. R. Sadler† and R. Y. Stanier

DEPARTMENT OF BACTERIOLOGY, UNIVERSITY OF CALIFORNIA, BERKELEY

Communicated by I. Michael Lerner, August 11, 1960

Our present knowledge of the physiology and biochemistry of green sulfur bacteria (genus *Chlorobium*) is largely based on the studies of van Niel[1] and of Larsen.[2] Their work showed that the members of this genus are strictly anaerobic, obligate photolithotrophs, which perform a typical bacterial photosynthesis using reduced inorganic sulfur compounds as electron donors. The physiology of these photosynthetic bacteria is accordingly similar to that of the purple sulfur bacteria, with one significant exception: all members of the latter group so far studied in pure culture can also use a variety of simple organic substrates for photosynthesis in the absence of a reduced inorganic sulfur compound. The possible participation of organic compounds in the photometabolism of green bacteria was examined in detail by Larsen,[2] using *Chlorobium thiosulfatophilum*. He found that none of the simple organic compounds commonly used as substrates for photosynthesis by purple sulfur bacteria could support the photosynthetic growth of *C. thiosulfatophilum*. However, manometric experiments suggested that resting cells could bring about a slow and limited transformation of organic acids anaerobically in the light. Only with propionate was a rapid metabolism observed. More detailed study showed that *C. thiosulfatophilum* performs a unique photosynthetic reaction with this substrate, grossly expressed by the equation:

$$CH_3CH_2COOH + CO_2 \xrightarrow{\text{light}} HOOCCH_2CH_2COOH. \quad (1)$$

A large part of the propionic acid metabolized is excreted into the medium as succinic acid. The physiological function of this reaction is obscure. Larsen[2] commented: "it can be safely concluded that the formation of succinic acid can hardly be considered as an important stage in the normal assimilatory process." However, the existence of the reaction revealed one very important fact: the failure of green bacteria to utilize organic compounds for growth cannot be attributed to the failure of such compounds to enter the cell.

Recently, Mechsner[3] has reported that a new species of *Chlorobium*, *C. chlorochromatii*, can grow well photosynthetically at the expense of peptone and malate. He confirmed, however, that the two species studied by Larsen, viz. *C. limicola* and *C. thiosulfatophilum*, are unable to utilize organic compounds for growth.

This paper reports certain new observations on the utilization of organic substrates by *C. limicola*.

Materials and Methods—Organisms and culture conditions: Several strains of *C. limicola* were isolated from local mud samples using the procedures described by Larsen[2] for enrichment and purification. Larsen's mineral medium,[2] which contains 0.024 *M* $NaHCO_3$ and 0.0042 *M* Na_2S, was used routinely for cultivation. In this medium, sulfide is the growth-limiting nutrient, as shown by preliminary experiments in which the sulfide content was varied, other components being kept constant. There is strict proportionality between cell yield and sulfide concentration

up to a sulfide concentration of 0.0073 *M*. In the experiments to be described Larsen's basal medium was used, but the concentrations of bicarbonate and of sulfide were sometimes varied; such variations will be indicated in the text. Stock cultures were maintained as stabs in Larsen's mineral medium solidified with 2 per cent agar, and covered with a layer of vaspar.

Procedures for growth experiments: Cultures were grown in completely filled, screw-capped bottles of 60 ml capacity, incubated at approximately 30°C in a light cabinet. Two methods for estimating the density of bacterial cultures were used: light scattering, measured with a Beckman DU spectrophotometer at 600 mμ, and determination of protein by the method of Lowry *et al.*[4] The simpler turbidimetric method cannot be used when cultures contain elemental sulfur, as is the case until the sulfide has been completely oxidized. Comparative measurements on sulfur-free cell suspensions showed that the two methods give strictly comparable results: a turbidity of 1.000 at 600 mμ corresponds to a protein content of 218 μg/ml.

In many experiments, the chlorophyll content of the culture was also determined, using a modification of the method of Cohen-Bazire *et al.*[5] for the determination of bacteriochlorophyll. A measured volume of the culture was centrifuged, the supernatant was discarded, and the packed cells were resuspended in 0.1 ml of water. After addition of 4.9 ml of acetone-methanol (7:2 v/v) the suspension was mixed and again centrifuged. The optical density of the supernatant extract was immediately measured at 665 mμ. The chlorophyll content was calculated using the specific absorption coefficient of 92.6 1/gm.cm for chlorobium chlorophyll-660 in acetone reported by Stanier and Smith.[6]

Measurements of radioactivity: The radioactivity of cells which had incorporated C^{14}-labeled acetate or bicarbonate was determined as follows: A suitable aliquot of culture was drawn through a Millipore filter and the cells on the filter washed with water. The filter pad was then removed, fastened onto an aluminum planchet with a thin layer of paste,[7] dried, and counted. For determining the intracellular distribution of radioactivity, washed cells were fractionated according to the procedure of Roberts *et al.*[8] All radioactive samples were counted in a Nuclear-Chicago gas flow counter.

Results.—Effect of organic substrates on the growth of C. limicola: In confirmation of earlier reports[1–3] we could not obtain photosynthetic growth of *C. limicola* at the expense of any of the organic compounds tested. However, we observed that certain organic compounds added at low concentration (0.005 *M*) to the basal medium of Larsen (in which growth is ultimately limited by the amount of sulfide furnished) caused detectable increases in growth yield and in final chlorophyll content. Results of a typical experiment are shown in Table 1. Small but significant increases in both growth yield and chlorophyll production were obtained with propionate, pyruvate, lactate, glucose, and glutamate, and a very marked increase with acetate. Accordingly, it was decided to investigate the acetate effect in greater detail.

The increment in growth yield obtainable with different quantities of acetate was first ascertained. Addition of increasing quantities of acetate to the basal medium systematically increased the cell yield and the chlorophyll content of cultures up to an acetate concentration of approximately 0.005 *M*. As the acetate

TABLE 1

EFFECT OF ORGANIC SUPPLEMENTS ON CELL YIELD AND CHLOROPHYLL CONTENT OF *C. limicola* GROWN PHOTOSYNTHETICALLY AT THE EXPENSE OF SULFIDE AND BICARBONATE

Organic supplement	Cell yield (Optical density at 600 mμ)	Chlorophyll content (mg/100 ml of culture)
None	0.420	0.316
Acetic acid	0.810	0.585
Propionic acid	0.505	0.356
Butyric acid	0.430	0.328
Pyruvic acid	0.530	0.430
Succinic acid	0.408	0.310
Malic acid	0.408	0.295
Lactic acid	0.450	0.360
Glucose	0.470	0.373
Glutamic acid	0.468	0.370

All organic compounds were added at a final concentration of 0.005 *M* to Larsen's basal medium. Organic acids were used as sodium salts.

concentration was further increased, both cell yield and chlorophyll content remained at the maximal value until the acetate concentration approached 0.1 *M;* at this level, both values began to decline. The maximal increase in cell yield and chlorophyll content represented an approximate doubling of the values obtainable without any addition of acetate.

The metabolic fate of acetate: Cultures were grown in the basal medium supplemented with 0.004 *M* $C^{14}H_3C^{14}OOH$ (90,000 counts/min/culture). Determinations of the radioactivity of the cells showed that 78 per cent of the acetate carbon was incorporated into cell material after 3 days of growth, and 93 per cent after 6 days.

In order to compare the gross intracellular distribution of carbon derived from acetate and from CO_2, the following experiment was performed. Three cultures were prepared, one containing $C^{14}O_2$ as the sole carbon source, one containing $C^{14}O_2$ and unlabeled acetate, and one containing acetate-C^{14} and unlabeled CO_2. The percentage incorporation of the labeled substrate and its distribution among the various cell fractions were determined in each case (Table 2). Under the conditions of this experiment acetate is almost entirely incorporated into cell material and the intracellular distribution of acetate carbon is grossly indistinguishable from that of CO_2 carbon. Furthermore, the gross intracellular distribution of carbon derived from CO_2 is not significantly altered when acetate is simultaneously assimilated. The magnitude of CO_2 incorporation appears to be slightly reduced by the simultaneous uptake of acetate, but the difference is barely significant.

TABLE 2

INTRACELLULAR DISTRIBUTION OF CARBON ASSIMILATED FROM CO_2 AND ACETATE BY *C. limicola*

	Nature of Carbon Source		
Radioactive substrate incorporated in cells (μmoles)	$C^{14}O_2$ 380	$C^{14}O_2$ and $C^{12}H_3C^{12}OOH$ 317	$C^{13}O_2$ and $C^{14}H_3C^{14}OOH$ 223
	Percentage of total counts in cells		
Cold TCA-soluble	2	2	3
Ethanol-soluble	32	34	36
Ethanol-ether-soluble	6	4	7
Hot TCA-soluble	9	8	6
Residual protein	34	38	35
Total recovery of counts in cells	83	86	87

Cells grown in Larsen's basal medium, furnished in all cases with 1440 μmoles of CO_2. When acetate was added, the amount was 240 μmoles.

Dependence of acetate assimilation on provision of sulfide and CO_2: *C. limicola* was grown in two series of bottles containing amounts of sodium sulfide ranging from 26 μmoles to 400 μmoles. One series contained no acetate and the other was supplemented with a fixed concentration of $C^{14}H_3C^{14}OOH$ (0.004 *M*). The growth yields for both series were determined, and the percentage incorporation of acetate carbon into the cells was also determined in the series supplemented with radioactive acetate. The data in Figure 1 show that the absolute magnitude of the increase in growth yield produced by acetate is proportional to sulfide concentration. As would accordingly be expected, the amount of acetate carbon assimilated by the cells is proportional to the amount of sulfide furnished (Fig. 2). For every mole of sulfide oxidized, about 0.75 mole of acetate is assimilated.

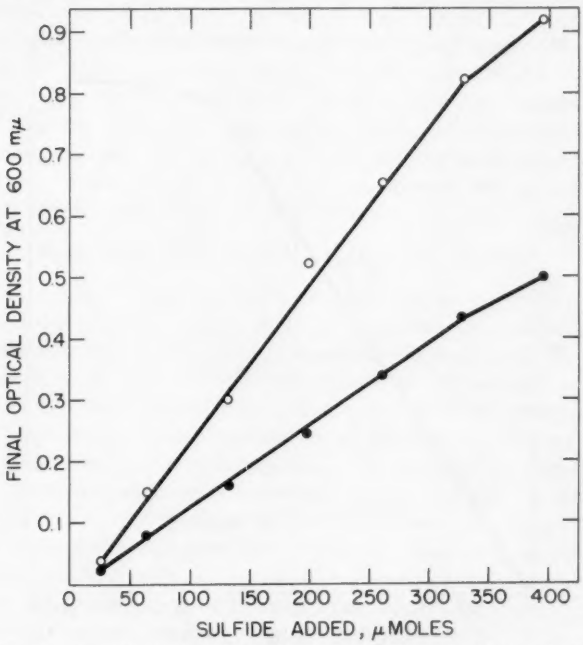


FIG. 1.—Growth yields of *C. limicola* as a function of sulfide concentration. Closed circles: Cells grown in basal medium alone. Open circles: Cells grown in basal medium supplemented with 0.004 *M* acetate.

The influence of CO_2 on acetate assimilation was examined by growing *C. limicola* in two series of bottles containing a fixed amount of sulfide (260 μmoles) and varying amounts of sodium bicarbonate (7.5 μmoles to 1500 μmoles). One series was supplemented with 52.5 μmoles of acetate-C^{14} and the other with 105 μmoles of acetate-C^{14}. After growth, the amount of acetate carbon incorporated into the cells was determined (Table 3).

With 260 μmoles of sulfide, the amount of bicarbonate which can theoretically be assimilated according to the gross photosynthetic equation:

$$H_2S + 2CO_2 + 2H_2O \rightarrow 2(CH_2O) + H_2SO_4 \quad (2)$$

is 520 μmoles. Accordingly, bicarbonate is in excess at the two highest levels employed. At these levels, there is a virtually quantitative assimilation of acetate carbon, as would be expected from the previously established quantitative dependence of acetate assimilation on sulfide concentration (Fig. 2). However, at levels where bicarbonate is the growth-limiting nutrient (375 μmoles and below), progressively less acetate is assimilated as the bicarbonate level is reduced. Within this region of bicarbonate limitation, the amount of acetate taken up by the cells is independent of the amount of acetate added but is strictly dependent on the amount of bicarbonate added. A small amount of acetate is assimilated in the absence of added bicarbonate; however, as the inoculum for this experiment was taken from a culture which contained excess bicarbonate, some carry-over of CO_2 no doubt

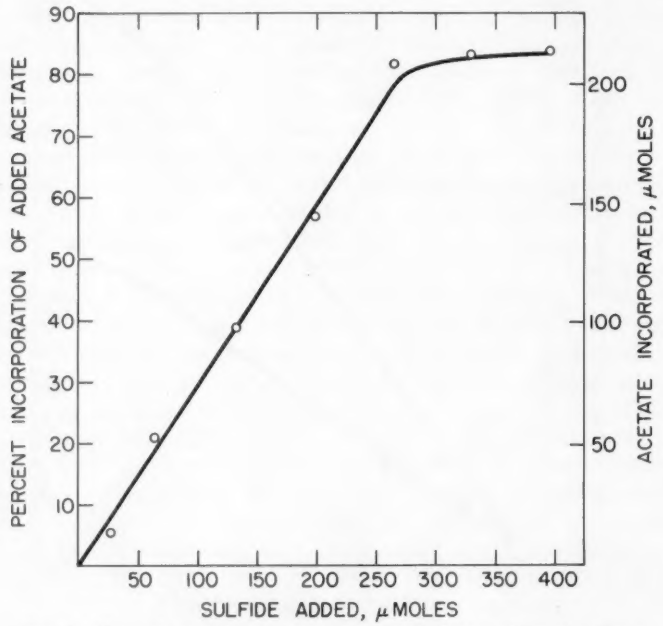


FIG. 2.—Dependence of acetate assimilation by *C. limicola* on the concentration of sulfide. Cells grown in basal medium, supplemented with the indicated amounts of sulfide and with a fixed amount of $C^{14}H_3C^{14}OOH$ (240 μmoles).

TABLE 3

INFLUENCE OF BICARBONATE CONCENTRATION ON ASSIMILATION OF ACETATE BY *C. limicola*

$NaHCO_3$ added to medium (μmoles)	Acetate Content of Medium	
	52.5 μmoles	105 μmoles
	Acetate incorporated in cells (μmoles)	
1500	48.8	96.5
750	47.8	103.0
375	49.3	87.0
75	31.0	26.3
37.5	23.1	22.0
7.5	13.1	13.6
None	7.9	11.5

occurred. This fact makes it impossible to determine from the data the absolute quantitative relationship between acetate assimilation and CO_2 assimilation.

Interpretation of the results: These experiments show that acetate can serve as a general source of cellular carbon during the photosynthetic growth of *Chlorobium limicola*. It can only do so, however, when the cells are simultaneously supplied with CO_2; furthermore, the amount of acetate which can be assimilated is proportional to the amount of CO_2 available. These experiments also show that acetate cannot provide the reducing power required for photosynthesis: its utilization by *Chlorobium* is strictly dependent on the provision of H_2S, and the amount of acetate which can be assimilated is proportional to the amount of H_2S available.

Thus, contrary to earlier suppositions,[1, 2] an organic compound can serve as a carbon source for *C. limicola;* however, the conditions under which it can be so utilized are severely restrictive. This is in sharp contrast to the manner of acetate utilization characteristics of purple bacteria. Anaerobically in the light, these organisms can grow photosynthetically using acetate as sole source of carbon and of reducing power.[9] The gross mechanism of this photosynthetic process has become reasonably clear as a result of recent work with *Rhodospirillum rubrum*.[10] Acetate undergoes a primary anaerobic oxidation-reduction. It is in part oxidized to CO_2, presumably via the tricarboxylic acid cycle, and the reducing power so made available is used for the reductive synthesis of a primary reserve material, poly-β-hydroxybutyric acid, from other acetate molecules:

$$9n \text{ acetate} \xrightarrow{\text{light}} 2n\ CO_2 + 4(C_4H_6O_2)_n + 6n\ H_2O. \qquad (3)$$

The function of light is to provide the needed ATP for the synthetic reactions.[9, 10] In *R. rubrum*, subsequent conversion of the primary reserve material to other cell materials (carbohydrate, protein, etc.) is rigorously dependent on the presence of CO_2.[10] However, even in a culture initially provided with acetate as sole carbon source, CO_2 will have been formed in the primary oxidation-reduction (see equation (3)) and will therefore be available for subsequent general cellular synthesis.

In the light of this work, a simple hypothesis can be offered to explain the much more restricted manner of acetate utilization by *Chlorobium*. If this organism lacks the enzymatic machinery required for the oxidation of acetate, it would be unable to use acetate either as a source of reducing power or as a source of CO_2; under these conditions, the use of acetate as a source of cellular carbon would be dependent on the provision of reducing power and of CO_2 from other sources. Such a double dependence is precisely what we have observed.

Confirmation of this hypothesis will require analysis on the biochemical level. Nevertheless, our present physiological studies have revealed an additional fact which lends some support to it. It was shown by Muller[11] that cultures of purple sulfur bacteria growing at the expense of acetate produce approximately 0.17 moles of CO_2 per mole of acetate metabolized. This value is slightly lower than the value predicted by equation (3), as would be expected if some of the CO_2 formed from acetate is subsequently fixed during the conversion of the primary reserve material to general cellular materials. In our experiments on the assimilation of acetate-C^{14} by *Chlorobium*, we have made several determinations of the fraction of the total radioactivity present in CO_2 after the complete utilization of acetate. This fraction was always very small, and its value was therefore difficult to deter-

mine with accuracy; however, the observed ratio of counts incorporated into cell material to counts evolved as CO_2 was of the order of 90:1. This corresponds to a molar ratio of acetate assimilated to CO_2 evolved of approximately 0.02, some ten per cent of the value which can be inferred from Muller's data on acetate metabolism by purple sulfur bacteria. It is therefore evident that *Chlorobium* has a much more limited ability to form CO_2 from acetate than do the purple sulfur bacteria.

* This work was supported by a grant from the National Science Foundation.

† Present address: Department of Biochemistry, University of Sydney, Australia.

[1] van Niel, C. B., *Arch. Mikrobiol.*, **3,** 1 (1931).

[2] Larsen, H., *Kgl. Norske Videnskabers Selskabs Skrifter*, 1 (1953).

[3] Mechsner, K., *Arch. Mikrobiol.*, **26,** 32 (1957).

[4] Lowry, O. H., N. J. Rosebrugh, A. L. Farr, and R. J. Randall, *J. Biol. Chem.*, **193,** 265 (1951).

[5] Cohen-Bazire, G., W. R. Sistrom, and R. Y. Stanier, *J. Cell. Comp. Physiol.*, **49,** 25 (1957).

[6] Stanier, R. Y., and J. H. G. Smith, *Biochim. Biophys. et Acta* (in press).

[7] Nielsen, E. S., *Nature*, **167,** 684 (1951).

[8] Roberts, R. B., D. B. Cowie, P. H. Abelson, E. T. Bolton, and R. J. Britten, *Carnegie Institution of Washington Publication* **607** (1955).

[9] Losada, M., A. V. Trebst, S. Ogata, and D. I. Arnon, *Nature*, **186,** 753 (1960).

[10] Stanier, R. Y., M. Doudoroff, R. Kunisawa, and R. Contopoulou, these PROCEEDINGS, **45,** 1246 (1959).

[11] Muller, F. M., *Arch. Mikrobiol.*, **4,** 131 (1933).

---

*ON THE ELECTRON-DONATING PROPERTIES OF INDOLES**

BY ALBERT SZENT-GYÖRGYI AND IRVIN ISENBERG

THE INSTITUTE FOR MUSCLE RESEARCH, MARINE BIOLOGICAL LABORATORY, WOODS HOLE, MASSACHUSETTS

Communicated August 13, 1960

It has been noted that indoles, such as 5-hydroxyl tryptamine (serotonin) form charge transfer complexes with riboflavin,[1] and that this charge transfer is a strong one, involving the formation of a primarily ionic complex and leading, under suitable conditions, to the formation of free radicals.[2] The orbital energies of indole, calculated by the Pullmans,[3,4] do not predict such strong electron donor properties. They show indole to be only a fair electron donor. That this is the case was shown experimentally by E. Fujimori,† who measured the charge transfer spectrum of indole with various classical acceptors, capable of π–π interaction, like trinitrobenzene, benzoquinone, chloranil, and bromanil.

It seemed desirable to find the cause of this discrepancy since various biologically important molecules, like serotonin, tryptophan, and growth hormones, are indole derivatives.

The electronic structure of indole, calculated by the Pullmans,[3] shows a high formal negative charge on carbon atom 3, and a smaller one on carbon 2. It seemed thus possible that the strong electron donor properties of indole were due to the high electron density on one or both of these C atoms. The π electrons of

these carbon atoms belong to the π electron pool of the highly conjugated molecule. So, it seemed possible that the high electron density makes these individual C atoms into good "local" donors, which could draw on the π electron pool of the whole molecule for the donated electron. As the work of Fujimori showed, this localized electron-donating property does not declare itself in a π–π interaction. However, it might declare itself in an interaction with a small, or "local," electron acceptor. Such an interaction would not demand the rigorous steric requirements on a donor-acceptor complex that a π–π interaction of two aromatic structures does. For this reason we have tested the interaction of indoles with iodine, the I_2 molecule being a relatively small molecule and a fair electron acceptor.

Charge transfer will occur only if there is overlap between the donor and the acceptor orbitals. At room temperature, in solution, thermal agitation may keep the two molecules apart. In order to study charge transfer, it may be necessary to force the two together, which can be done in various ways. One can, for instance, dissolve both substances in water and freeze the solution, whereby the solvent crystallizes out leaving the solutes in close proximity. In case of nonpolar substances, one can dissolve these separately in a volatile solvent like chloroform, then moisten a filter paper with the solution of the one, say the donor. After the solvent has evaporated, one can moisten the paper with the solution of the acceptor. After evaporation of the solvent, the two solutes will be left in close proximity. One can also evaporate a solution of the mixture of both components. In many cases it may suffice to lower the temperature of the mixture of the two substances in order to induce the desired complex formation. Another method consists of allowing a solution of the donor in question to dry on filter paper, then exposing the paper to the action of I_2 vapor.[5]

An 0.1 *M* solution of I_2 in chloroform is purple. If this solution is mixed with an 0.1 *M* solution of indole, the solution remains purple showing that little or no charge transfer complex is formed. However, if the solution is placed on filter paper and the solvent allowed to evaporate, a black spot is left behind. Similarly, if the solution is cooled in a dry-ice cooling mixture a voluminous black precipitate is formed which dissolves again on warming. The precipitate may be separated by centrifugation at $-20°$C. When the precipitate is pressed out between sheets of filter paper and quickly air dried, a black powder is obtained. Such a powder was placed into the cavity of an electron spin resonance apparatus and a high and narrow signal was obtained, showing 10^{19} spins per gram.‡ The signal showed no hyperfine structure and had a g value of 2.001, close to that of a free electron. This indicates that a strong charge transfer has occurred from indole to iodine. A similar black precipitate was obtained with skatole, indicating that the nitrogen atom was not involved in the reaction. Indole acetic acid and serotonin gave similar reactions. In these molecules the side chain does not appreciably alter the charge of the C atom at position 3 and 2. However, when the high negative formal charges of C2 and C3 were abolished by making these part of an additional conjugation, as is the case in carbazole, no black complex was formed, although the NH remained free. This supports the assumption that the C atoms 3, or 2, (or both), were involved in the formation of the charge transfer complex. Free pyrrole has a pronounced negative charge on both these C atoms and gives a similar reaction as in indole.§

These reactions deserve attention because it seems possible that the described high electron-donating activity in indole plays a role in the reactions of proteins, since the C3 atom is conjugated by a hyperconjugated CH_2 to the peptide link, which may form through the H-bridges a continuous electronic system with the other peptide links of the same molecule. It may play a role also in the biological activity of indoleacetic acid and serotonin.

The black indole-I_2 complex is, in all probability, in many ways analogous to the black iodine complexes formed by molecules with extensive conjugate systems, like perylene or violanthrene, which also give strong electron spin resonance signals.[6, 7] This suggests that while extensive conjugated systems may tend to donate electrons to a local acceptor in any case, smaller and simpler molecules with conjugated systems can do so only if they contain highly charged carbon atoms acting as "local donors." Though "local" the donating atom may draw on the π pool of the molecule for the electrons to be donated.

Note added in proof: An elementary analysis of the indole-iodine complex has been made by the Schwarzkopf Microanalytical Laboratory and showed 27 *N* atoms for 23 I_2 molecules suggesting that the complex contained one I_2 per indole.

* This research was supported by a grant from the Commonwealth Fund, Grant No. H-2042(C3) of the National Heart Institute, a grant from the National Science Foundation, and Grant No. 1 from the Quartermaster Research & Engineering Command, U. S. Army (Natick).

† Fujimori, E., unpublished work, see quotation 4, page 63.

‡ The substance probably contained some moisture, due to the condensation of water vapor at the low temperature employed.

§ Due to the volatility of pyrrole, it is advisable to allow an 0.1 *M* solution in chloroform, mixed with equal or double volumes of 0.1 *M* I_2 in chloroform, to dry on filter paper instead of applying the two substances one by one.

[1] Isenberg, I., and A. Szent-Györgyi, these PROCEEDINGS, **44,** 857 (1958).

[2] Isenberg, I., A. Szent-Györgyi, and S. L. Baird, Jr., these PROCEEDINGS, (in press).

[3] Pullman, B., and A. Pullman, in *Les théories électroniques de la Chimie Organique,* (Paris: Masson, 1952).

[4] Szent-Györgyi, A., in *Submolecular Biology,* (New York: Academic Press, 1960).

[5] Hegyeli, A., *Biol. Bull.,* (in press).

[6] Akamatu, H., H. Inokuchi, and Y. Matsunaga, *Nature,* **173,** 168 (1954).

[7] Matsunaga, Y., *J. Chem. Phys.,* **30,** 855 (1959).

THE FUNCTION OF RNA IN T2-INFECTED BACTERIA

BY ELLIOT VOLKIN

BIOLOGY DIVISION, OAK RIDGE NATIONAL LABORATORY,* OAK RIDGE, TENNESSEE

Communicated by Alexander Hollaender, August 26, 1960

When *Escherichia coli* B are infected with virulent bacteriophages, no net synthesis of RNA can be demonstrated.[1–6] Nevertheless, an active metabolism of RNA has been shown to exist.[6–8] Employing either P^{32}- or C^{14}-labeled nucleic acid precursors, we determined[6, 8, 9] that some of the host RNA is concomitantly synthesized and degraded (or utilized). This metabolic turnover of RNA is most extensive in the early stages after infection, exceeding the incorporation rate of isotope into DNA the first few minutes after infection.[10] Unlike the RNA in

uninfected *E. coli* B whose constituent nucleotides become equally labeled with isotopic precursors,[10, 11] the RNA mononucleotides of the phage-infected cell become unequally labeled; the RNA labeling pattern is distinct for each type of infecting phage and simulates the composition of the particular phage DNA.[6, 8, 9] Experiments involving a centrifugal separation of the host's subcellular components reveal a marked heterogeneity in the labeling behavior of the RNA from different subcellular fractions.[12] Recently, Nomura *et al.*[13] using more refined methods of isolation of bacterial cell constituents, have succeeded in purifying to a great extent the RNA that is metabolically most active after T2 infection. These and other data[14] are consistent with our proposal[6, 12] that the synthesis of a new and specific species of RNA, relatively small in amount with respect to the total host RNA, is induced by phage infection.

Some evidence from other laboratories has also implicated an RNA involvement in phage production. Jeener has shown[15] that the ultraviolet induction of phage protein produced by a lysogenic strain of *B. megatherium* is prevented by treatment of the bacteria with RNase or the presence of various metabolic antagonists to RNA synthesis. The reversal by uracil of inhibition of T2 replication by 6-azauracil is interpreted by Pardee[16] to be a function of some early RNA synthesis.

In seeking for a means to assess a biologic function for RNA in phage growth, we were guided by the observation that some synthesis of RNA unique to the phage-infected cell can occur during inhibition of protein synthesis; that is, by administration of chloramphenicol[14, 17] or by deprivation of the necessary amino acid with an amino acid-requiring strain of bacteria.[18] Thus, with the use of certain amino acid-requiring mutants of coli B, RNA synthesis may proceed in the absence of protein (and DNA) synthesis. It appeared feasible to regulate either RNA or protein synthesis if a double mutant, requiring a nucleic acid precursor as well as an amino acid for growth, was employed. The mutant B94, whose properties have been described by Gollob and Gots,[19] requires both adenine and arginine for growth and has now been found to need these metabolites for T2 production. Although deoxyadenosine will not substitute for adenine in supporting phage growth, it is possible to preincubate the infected cells in the presence of adenine and thereby allow subsequent phage synthesis in an adenine-free medium containing deoxyadenosine and arginine. The data obtained lead us to the provisional conclusion that the synthesis of specific RNA(s) during the preincubation in adenine permits the ensuing yield of bacteriophage.

Materials and Methods.—Chemicals: Adenine and deoxyadenosine were obtained from the California Corporation for Biochemical Research and were their chromatographically homogeneous grades, labeled CfP. L-(+)-arginine monohydrochloride was obtained from Eastman Kodak Company. These compounds were used without further purification and, unless otherwise stated, in a final concentration of 40 μg/ml. The analogue 6-azauracil, B grade, obtained from the California Corporation for Biochemical Research, was used in a final concentration of 250 μg/ml. A stock of chloramphenicol (Parke, Davis Company, Detroit, Michigan), in a concentration of 1 gm/100 ml in 50% alcohol, was maintained at refrigerated temperatures. P^{32} was obtained from the Oak Ridge National Laboratory as orthophosphoric acid in dilute HCl and was neutralized with $NaHCO_3$ before use.

Biologic methods: Unless otherwise specified, the minimal medium used consisted of a glucose (0.4%)-salts mixture that is low in phosphorus and buffered at pH 8.0 with tris.[6] The source and maintenance of the stocks of *E. coli* B, *E. coli* B94,† and T2r$^+$, and the purification of the bacteriophage have been described elsewhere.[11, 20] A 1 *M* phosphate buffer, pH 5.9, consisting of 9 parts of KH_2PO_4 to one part of Na_2HPO_4 was used in many experiments as an aid to maintaining the integrity of phage-infected cells, as described in the Results section.

Biologic assays for free phage and infected bacteria were carried out according to standard procedures described by Adams.[21] *E. coli* B in active logarithmic growth were used as plating bacteria and serial dilutions were always carried out with peptone broth.[22] Chloroform was used for prematurely lysing infected bacteria and liberating intracellular phage. In general, 2-ml quantities of bacterial suspension were shaken briefly with 0.1 ml $CHCl_3$ and allowed to stand 5 min, after which the aqueous phase was assayed. Liquid cultures for bacterial growth or phage-infected cells were incubated at 37°C using rather vigorous aeration. Multiplicities of about 8 phage particles per bacterium were used.

Chemical methods: The method of chemical fractionation for the recovery and quantitation of the nucleic acids has been described.[6, 23] In brief, the method consists in the removal of perchloric acid soluble and lipid materials, followed by the extraction of nucleic acids by hot 2 *M* NaCl. Ribonucleic acid was converted to mononucleotides by mild alkali (0.1 *M*, NaOH, 37°, 16–18 hr), and DNA was precipitated by the addition of perchloric acid (to 0.4 *M*); after these treatments, it is necessary to centrifuge the DNA at relatively high speeds, e.g., 20,000 × *g* for 20 min, for complete sedimentation. The RNA mononucleotides were separated by ion-exchange chromatography from Dowex-2 (either 2× or 8×), chloride form, using dilute concentrations of HCl (0.001 *N* to 0.007 *N*) as the eluent.[24] Methods for measuring ultraviolet light absorption,[6] phosphorus content,[25] and radioactivity[6] have been adequately described.

A procedure was developed for estimating the radioactivity of the proximal 5′-phosphorus of the nucleotides in the free (acid soluble) nucleotide pool. It was discovered[26] that our preparations of snake venom diesterase, freed of 5-nucleotidase activity, contain a highly potent apyrase activity, whereby nucleoside di- and triphosphates are quickly and quantitatively converted to the 5′-monophosphates. The use of such an enzyme with radioactive nucleotides permits a rapid estimate of the radioactivities associated with the proximal 5′-phosphate, the putative precursor of nucleic acid phosphorus. A cold 0.4 *M* $HClO_4$ extract (200 ml) was stirred for 30 min with 1 gm powdered Darco G-60, which had been previously washed with 2 *N* HCl and substantial quantities of water. The charcoal was filtered through a sintered glass funnel of medium porosity and washed with water, after which the nucleotide material was eluted with ammoniacal alcohol (15 ml conc. NH_4OH in 50% alcohol). About 80% recovery of total nucleotides is obtained by this procedure. After evaporation of the eluent, the solids were resuspended in 10 ml of 0.05 *N* NH_3-NH_4Cl (pH 8.7). Snake venom diesterase was added and the solution incubated at 37°C until no further phosphorus was liberated as inorganic orthophosphate.[27] The material was then subjected to ion-exchange analysis, using Dowex-2 (8×, chloride form and chloride elution).[24] Since the quantities of nucleotides were minimal in these experiments, it was

necessary to add a mixture of authentic 5′-nucleotides as carrier for identifying the position and relative purity of the radioactive nucleotides; the 5′-deoxynucleotides of cytosine, adenine and guanine and 5′-uridylic acid were used as markers for this purpose. This procedure, therefore, allows a valid measure of the relative total counts contained in each nucleotide and does not permit an estimate of specific activities.

Results.—Nutritional requirements for T2 production by B94: We wished to establish whether the double requirement for adenine and arginine was necessary not only for growth of the bacterium but also for T2 production. It is known that some degradation of host DNA occurs after phage infection, and indeed host DNA is largely converted to phage DNA.[28] Since the host may contain deoxynucleoside phosphorylases, deoxynucleosidases, or nucleoside transdeoxyribosidases,[29] it is conceivable that, by the combined action of DNA degradative enzymes with either of the latter enzymes, the phage-infected cell could supply the required adenine by way of its DNA. In this respect it was pertinent to determine whether the deoxynucleoside of adenine could substitute for the free base in the phage-infected cell. Log-phase bacteria were washed with minimal medium, resuspended in the various media given in Table 1 and followed with respect to growth and

TABLE 1

GROWTH AND T2 PRODUCTION BY B94

Minimal medium plus	Noninfected (turbidity increase)	T2r⁺-infected (100 min, CHCl₃lysate), T2/B*
. . .	0	0.02
Ade.	0	0.03
Arg.	0	0.02
dAdo. + Arg.	0	0.02
Ade. + Arg.	++++ (doubling time ~60 min)	198

Bacteria were grown to a titer of about 4×10^8/ml in minimal medium fortified with adenine and arginine, then centrifuged, washed twice with minimal medium, resuspended and divided equally among the media shown in the table. The multiplicity of phage infection was 8 and no significant loss of infective centers occurred during the incubation period.

* T2/B = Intracellular phage per infected bacterium. Yield values (T2/B) of 0.05 or less, reported throughout this paper, are of the same order of magnitude as the value for unadsorbed phage, and they are in all probability due to the latter, rather than intracellular phage.

The following abbreviations are used in all tables and figures: Ade. = adenine, dAdo. = deoxyadenosine, Arg. = arginine.

ability to support T2 synthesis. The results show that both adenine and arginine are required for T2 production in B94. It should be noted that the combination of deoxyadenosine plus arginine will neither support phage production nor bacterial growth. It thus appears that the mutant does not have a mechanism for converting the deoxynucleoside to adenine, or that such a mechanism is inoperative under these cultural conditions.

P^{32} incorporation into RNA and DNA: The synthesis of RNA and DNA in phage-infected B94, cultured in minimal medium containing the various metabolites described above, was assessed by measuring their extent of incorporation of radioactive phosphorus. Washed cells were suspended in the different media, infected with T2r⁺, and given 5 μc/ml of P^{32} two min after infection. Samples were fractionated at 20 and 50 min after infection. The RNA values given in Table 2 represent the acid soluble material after alkaline hydrolysis of the total nucleic acid, without further purification by ion-exchange chromatography. As a result, the RNA counts as well as those for DNA ($HClO_4$ precipitable after alkaline hy-

TABLE 2

P^{32} INCORPORATION INTO RNA AND DNA OF T2-INFECTED B94

Minimal medium plus	Intracellular T2 product (T2/B) at 75 min	Activity at time: 20 min RNA	20 min DNA	50 min RNA	50 min DNA
		cts/sec/260-mμ unit			
	0.05	15	12	62	33
Arg.	0.05	42	8	44	18
dAdo. + Arg.	0.05	62	23	90	44
Ade.	0.05	196	44	418	89
Ade. + Arg.	150	344	1,150	827	6,520

drolysis) include some P^{32} containing contaminants and tend to be slightly higher than the true values. It is clear, however, that P^{32} incorporation into RNA is significantly greater only where adenine is present compared to the other three conditions where no intracellular phage is produced. The RNA labeling in the presence of adenine alone is about half that which takes place in the complete medium (adenine plus arginine) that permits phage synthesis, whereas DNA is only slightly labeled by comparison. The RNA synthesized during incubation in adenine does not undergo measurable turnover in this medium, inasmuch as no radioactivity is lost when P^{32} is replaced with nonradioactive phosphate, a situation similar to that existing in T2-infected *E. coli* B inhibited by chloramphenicol.[17]

Phage production in deoxyadenosine plus arginine after preincubation in adenine: On the basis of these results, experiments were designed to test for biologic activity that might reside in the RNA accumulated during the incubation of the T2-infected cell in adenine alone. A biologic specificity inherent in the labeled RNA is indicated by the typical RNA mononucleotide specific activity ratios obtained.[6, 8, 11] Thus, it seemed reasonable to suspect that such activity, if present, may be retained when the cells are deprived of the limiting RNA precursoral source (adenine) but permitted to synthesize protein and DNA by introducing arginine and deoxyadenosine.

The mutant was grown in minimal medium fortified with arginine and adenine to a titer of about 2×10^8 cells/ml, washed twice with minimal medium, then suspended in minimal medium containing adenine but no arginine. After five min incubation, the bacteria were infected with $T2r^+$ and incubation continued in this medium for one hour, during which time no intracellular phage was produced. The infected bacteria were then centrifuged, washed twice with minimal medium, incubated in minimal medium for 10 min, and given arginine and deoxyadenosine. It will be recalled that this medium containing deoxyadenosine and arginine, but not adenine, is incapable by itself of supporting T2 synthesis. Figure 1 (curve A) reveals that T2 production will now take place in the latter medium after preincubation in adenine. Under these conditions, the yields are limited to about 5–10 phage particles per infected bacterium, the full yield expressing itself within about 35 min, although mass lysis has not occurred at this time. It can be seen, however, that the addition of adenine at time 32 min (curve B) results in a further increase in phage production, demonstrating that the cell's potential for producing phage has not been damaged by these treatments. Curve C represents the normal rate of T2 production in medium complete with adenine and arginine, without having undergone preincubation in adenine alone. The yield of 5–10 phage particles per infected cell is not significantly altered by adding

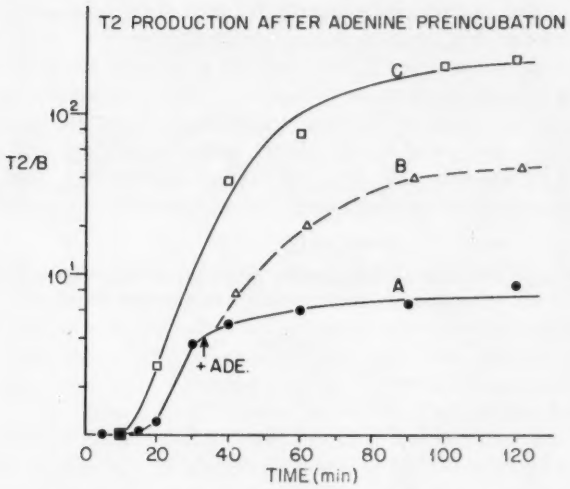


FIG. 1.—Curve A = Transferred to medium containing deoxyadenosine plus arginine after preincubation in adenine. Curve B = Adenine added to A at 32 min. Curve C = Complete system (adenine plus arginine) without adenine preincubation. For experimental details see text.

to the deoxyadenosine:arginine either casamino acids (100 μg/ml) or a mixture of the other three deoxynucleosides (40 μg/ml each), or a combination of both.

The limiting factor in phage yield, therefore, is not some nutritional requirement involved in protein or DNA synthesis but is some cell constituent requiring adenine as its precursor. That this constituent is not endogeneous, free adenine itself is strongly suggested by the fact that infected bacteria cultured in adenine alone, then washed twice, may be given an intermediate incubation in unfortified minimal medium for at least an hour, without significant loss in ability to produce phage when subsequently placed in medium containing deoxyadenosine and arginine.

Preincubation of infected bacteria in minimal medium, in medium containing deoxyadenosine alone, or in medium fortified with arginine alone will not support a later synthesis of intracellular phage in culture containing deoxyadenosine plus arginine, although no significant loss of infected bacteria occurs during these manipulations.

It is known that the synthesis of some protein early after infection is necessary to T2 DNA synthesis[30–32] and, consequently, for the formation of T2 progeny. The substrates utilized for T2 DNA synthesis are bacterial DNA as well as metabolites assimilated from the medium,[28] the latter representing a major precursoral source during the state of lysis inhibition with the r^+ phages. It was of interest, therefore, to examine the order of events with respect to protein and DNA syntheses required to yield viable phage progeny after infected cells had undergone a preincubation in adenine alone. Thus it might have been anticipated that such cells given arginine only would produce nearly a full complement of intracellular viral progeny, since the host DNA can serve as a T2 DNA source. On the other

hand, preincubated bacteria subsequently fed only deoxyadenosine would be unproductive, since protein synthesis would remain precluded.

After a 60-min incubation of T2-infected B94 in adenine alone, the washed cells were placed in minimal medium containing either deoxyadenosine alone or arginine alone. Intracellular phage production was assayed during an hour's incubation, then the bacteria were washed and placed in media fortified with the converse metabolite (i.e., arginine and deoxyadenosine, respectively), and intracellular phage again measured during the course of an additional hour's incubation (Fig. 2).

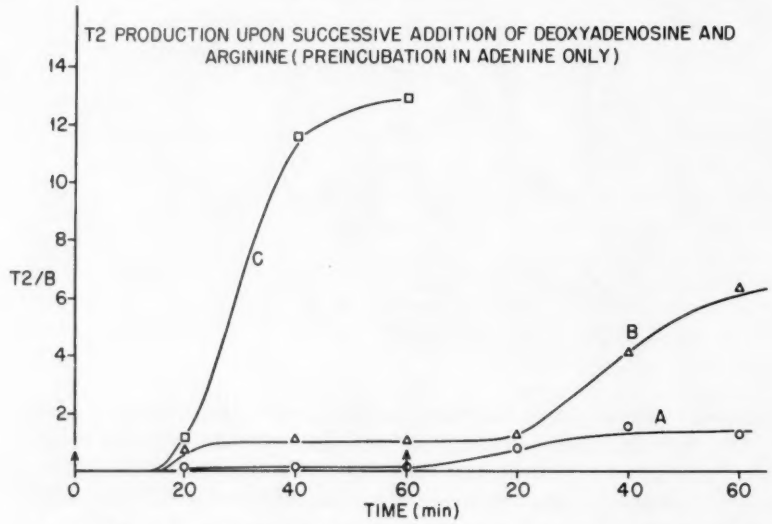


FIG. 2.—Infected bacteria were incubated for 60 min in minimal medium containing adenine only, during which time no intracellular phage was produced (not shown on the figure). The bacteria were then washed free of adenine and A = transferred to deoxyadenosine only for 60 min, then to arginine only for 60 min; B = transferred to arginine only for 60 min, then to deoxyadenosine only for 60 min; C = transferred to medium containing deoxyadenosine plus arginine.

Curve A of the figure demonstrates that deoxyadenosine alone, after adenine culture, will not support T2 formation, illustrating the need for some protein synthesis. Arginine alone, after adenine incubation, supports a limited extent of phage synthesis (curve B, first 60 min). It is noteworthy that the limited ability of arginine to support phage production is maintained after an hour's culturing in deoxyadenosine and removal of the deoxynucleoside (curve A, second 60 min). The latter result is likewise not in accord with the idea that a residual pool of free adenine is responsible for phage production. In view of the infected cells' ability to convert host DNA to phage DNA, it is not only surprising that the yield in arginine alone is so low (about 1 T2/bacterium), but even more remarkable is the fact that having removed the arginine, a large increase in phage yield (to about 6 T2/bacterium) occurs by feeding deoxyadenosine alone (curve B, second 60 min). This latter value is about half that obtained when the deoxynucleoside and arginine are present together (curve C). Thus it appears that the synthesis of some

T2 precursor protein occurs during the arginine incubation that can be utilized at a later time when DNA synthesis is more fully realized.

These data are interpreted as indicating the following sequence of events in the synthesis of T2 bacteriophage. There is a primary necessity for the synthesis of RNA (or possibly some other adenine-containing component, other than DNA); protein synthesis must follow for phage production, but for full phage productivity an exogenous source of DNA precursoral material must be present.

Effect of phosphate and pH lowering on the stability of infected cells: We have been able to increase the yield of bacteriophage after adenine preincubation by a change in the composition of the medium. Our minimal medium is low in phosphate, being buffered with tris at pH 8.0. If labeled, infected cells are centrifuged and washed with this medium, although there is only a small loss in cell count and total RNA, there is considerable loss (about 50%) of labeled RNA, indicating that the newly synthesized RNA preferentially moves out of the bacteria during these operations. This can be prevented by adding phosphate buffer before centrifuging both to increase the phosphate concentration and lower the pH. One fifteenth volume of 1 *M* phosphate buffer at pH 5.9 is added before centrifugation bringing the final pH to 6.4–6.5. The centrifuged cells, washed and resuspended in this phosphate medium, retain most of the labeled RNA, and phage yields are increased about threefold, as shown in Table 3. Phage yields now average

TABLE 3

EFFECT OF PHOSPHATE BUFFER ON T2 YIELD

Initial medium (60 min incubation)	T2/B at 60 min	Medium after transfer	T2/B at 60 min
Minimal (pH 8) + Ade.	0.02	Minimal (pH 8) + dAdo. + Arg.	7.4
Minimal (pH 8) + Ade.	0.02	Min:P + dAdo. + Arg.	28.7
Minimal (pH 8) + Ade.	0.02	Min:P + Ade. + Arg.	110
Min:P* + Ade.	0.02	Min:P + dAdo. + Arg.	26.6

Min:P* = Minimal medium containing 0.067 *M* phosphate buffer. The pH of the medium is 6.4.

20–30 intracellular particles per bacterium in the deoxyadenosine:arginine medium after adenine preincubations of an hour or more. The rest of the experiments to be reported were carried out with this technique.

Increased yield in complete medium: A consequence of forming some necessary intracellular constituent, presumably RNA, during incubation of adenine would be cells primed to replicate phage progeny, when placed in complete medium, at a rate faster than those in complete medium but not having undergone such preincubation in adenine. That such is indeed the case is shown in Table 4. In this experiment, one sample of bacteria, grown in medium containing adenine

TABLE 4

INCREASED YIELD OF T2 IN COMPLETE MEDIUM AFTER PREINCUBATION IN ADENINE MEDIUM

Minutes in complete medium	T2/B*	
	Control (no preincubation)	Experimental (after 75′ adenine incubation)
10	0.1	0.6
20	2	25
30	37	64
40	66	106
60	136	153
120	269	306

* The number of infected bacteria (1.90×10^8/ml) was the same in both samples.

and arginine, was directly infected with T2r$^+$ and intracellular phage was assayed at regular intervals. A parallel bacterial sample was washed twice with minimal medium, infected in the presence of adenine only, allowed to incubate 75 min (during which time no intracellular phage was detectable), then given arginine and intracellular phage measured during the same time intervals. It can be seen that the yield at 20 min in the latter culture is 25 compared to two in the control culture. It can also be seen that an increased yield of some 20–30 phage particles per bacterium at all times after 20 min is maintained by the culture that received the preincubation in adenine. This increment is the same order of magnitude as the total phage yield obtained when cells are subsequently given deoxyadenosine: arginine instead. Under the latter conditions, however, the maximum yield is obtained in 30 to 35 min incubation.

Correlation between RNA synthesis and phage production: To explore further whether a real relationship exists between RNA synthesis and subsequent phage production, we designed an experiment to test for a possible quantitative correlation between the extent of RNA synthesized during incubation in adenine and the corresponding extent of phage production upon placing the cells in medium containing deoxyadenosine:arginine. Washed cells were suspended in medium containing adenine alone, infected with T2r$^+$, one min later given $P^{32}O_4$, then at various intervals after incubation in adenine centrifuged and washed again. One aliquot of the resuspended cells was used for determining the amount of incorporation of radioactive phosphorus into RNA and DNA; another aliquot was used for measuring the maximum yield of intracellular phage produced on further incubation in deoxyadenosine:arginine. Figure 3 shows the data from two such experiments, where it may be observed that the curve for RNA synthesis in adenine is in reasonably good fit with the curve for ultimate phage yield while the curve for P^{32} incorporation into DNA is distinctively dissimilar.

RNA specificity: If uninfected B94 cells are incubated in adenine alone over a comparable period of time, then placed in deoxyadenosine:arginine and immediately infected with T2r$^+$, *no* intracellular phage are produced. This indicates that some specific component, necessary for subsequent phage production, is produced only in the phage-infected bacteria during the adenine preincubation.

Isotope studies were carried out in order to distinguish which unique adenine-containing component (RNA, DNA or free nucleotides) might be formed only in bacteria infected by phage. Equal aliquots of washed bacteria were suspended in minimal media containing adenine alone. In one case, the bacteria were infected with T2r$^+$ and given $P^{32}O_4$ within a minute after infection; $P^{32}O_4$ was added to the other sample, in which no phage was present, at the same time. Samples were removed at three time intervals after isotope addition and analyzed for total radioactivity in DNA and RNA, as well as for radioactivities associated with the individual mononucleotides of the RNA. In addition, the nucleotide material in the acid-soluble fraction was recovered, totally converted to mononucleotides and chromatographed to estimate the relative total radioactivity in the four major 5′-mononucleotides (see *Materials and Methods*, section on *Chemical methods*). Table 5 shows that at all times during incubation in adenine alone the extent of P^{32} incorporation into RNA and the proximal phosphorus of the free nucleotide pool is practically the same, and DNA somewhat higher at the earlier times for

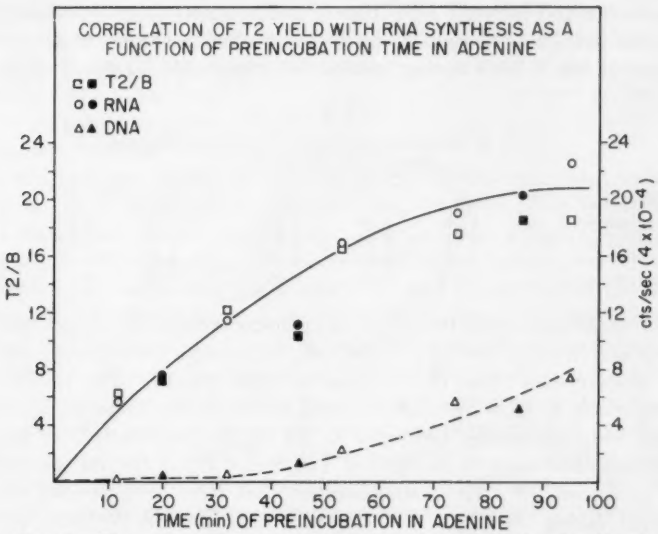

CORRELATION OF T2 YIELD WITH RNA SYNTHESIS AS A FUNCTION OF PREINCUBATION TIME IN ADENINE
T2/B
RNA
DNA
T2/B
cts/sec (4 x 10⁻⁴)
TIME (min) OF PREINCUBATION IN ADENINE

FIG. 3.—The scale representing total radioactivity (right hand ordinate) was arbitrarily chosen so that these values could be more easily compared with the values for intracellular phage (left hand ordinate). The points for T2/B correspond to the maximum phage yield obtained after transfer to arginine: deoxyadenosine, incubating 60 min and lysing with chloroform. The open and closed symbols represent points from two independent experiments.

TABLE 5

P^{32} INCORPORATION DURING INCUBATION IN ADENINE

Time, min	Total activity in RNA, cts/sec		Total activity in acid-soluble nucleotide proximal P, cts/sec		Total activity in DNA, cts/sec	
	Uninfected	T2-infected	Uninfected	T2-infected	Uninfected	T2-infected
20	10,000	10,800	9,800	11,000	2,300	420
45	16,600	16,300	15,100	15,700	3,700	2,030
85	27,300	27,300	20,100	22,200	7,020	7,150

TABLE 6

RATIOS OF TOTAL P^{32} IN PROXIMAL P OF ACID-SOLUBLE NUCLEOTIDES

Nucleotide	Uninfected			T2-infected		
	20′	45′	85′	20′	45′	85′
Cytidylic acid	1.0	1.0	1.0	1.0	1.0	1.0
Adenylic acid	2.9	3.6	3.0	3.5	2.2	2.7
Uridylic acid	9.2	7.9	10.1	7.4	7.9	6.2
Guanylic acid	3.3	4.6	4.7	4.6	6.0	4.6

uninfected bacteria relative to T2-infected bacteria. Likewise, Table 6 reveals that there are no regular or large differences in the relative labeling pattern of the innermost phosphorus of the free nucleotides at any time with respect to uninfected and T2-infected bacteria. It is clear, however, that there are large differences in radioactivity among the four nucleotides in both cases, where uridylic acid is generally considerably higher, cytidylic acid considerably lower, than the two purine nucleotides.

Thus, although the extent of the free nucleotide pool and the RNA formation in uninfected cells is similar to the T2-infected cells during incubation in adenine, the uninfected bacteria are incapable of producing phage when given arginine:

deoxyadenosine and infected with T2. The one outstanding difference detected between the two kinds of cells is the mode of labeling of the constituent mononucleotides of the RNA's during incubation in adenine. Table 7 demonstrates

TABLE 7

RNA MONONUCLEOTIDE SPECIFIC ACTIVITY RATIOS
(INCUBATION IN ADENINE)

Nucleotide	Uninfected			T2-infected		
	20′	45′	85′	20′	45′	85′
Cytidylic acid	1.00	1.00	1.00	1.00	1.00	1.00
Adenylic acid	0.97	0.95	0.89	1.60	1.54	1.27
Uridylic acid	1.11	0.93	0.88	1.66	1.43	1.27
Guanylic acid	0.94	0.91	0.97	1.00	0.86	0.83

that while in the uninfected B94 all RNA mononucleotides are approximately equal in specific activities, in the T2-infected cells the specific activities of adenylic and uridylic acids exceed those of the other two nucleotides. The specific activity ratios, especially at early incubation times, resemble the composition ratio of the analogous deoxynucleotides contained in T2 DNA; this result is in accord with previous data obtained with the RNA of T2-infected *E. coli* B in our laboratory,[6, 8, 9] as well as others.[13, 14] The results suggest that RNA of a specific structure is synthesized during the phase of incubation when protein synthesis is inhibited (adenine alone), and this RNA can exhibit some biologic function necessary for phage production when conditions allow protein and DNA synthesis but no further RNA formation (deoxyadenosine:arginine, but no adenine).

Chloramphenicol experiments: It may be surmised that during adenine incubation, in spite of the absence of arginine, there may remain a low level of protein synthesis that may account for the observed RNA synthesis. That protein synthesis itself may indeed be responsible for later phage formation in medium containing deoxyadenosine:arginine is ruled out by the negative response following arginine preincubation. In order to more fully inhibit protein synthesis, chloramphenicol (25 μg/ml) was added to the infected bacteria cultured in adenine alone, then the bacteria tested for intracellular phage productivity in both complete medium (adenine:arginine) and in deoxyadenosine:arginine. The results of these experiments reveal that, although there is some time delay in phage synthesis in both media, the ultimate phage yields are quite similar to those generally obtained without the addition of chloramphenicol to the adenine culture. If protein synthesis takes place during the adenine preincubation, then it is of a type that escapes the combined inhibitory actions of chloramphenicol inhibition and arginine deprivation.

6-azauracil experiments: In conformation with the results of Pardee and Prestidge,[16] we have found that 6-azauracil completely inhibits T2-bacteriophage formation in B94 cultured in adenine plus arginine. A further test of 6-azauracil was made in the B94 system by first preincubating T2-infected cells in adenine alone for 60 min, then adding the analogue to the subsequent culture media containing either arginine plus deoxyadenosine or arginine plus adenine. The infected bacteria cultured in arginine:deoxyadenosine containing 6-azauracil produced the same yield of phage (T2/B = 15 to 20) as those bacteria incubated in arginine:deoxyadenosine without 6-azauracil. This result is in accord with the concept that a functional pool of RNA is synthesized during the adenine preincubation and can

act in the presence of an inhibitor to further RNA synthesis. Bacteria subsequently cultured in complete medium (adenine plus arginine) containing 6-azauracil again produced only 15–20 phages per infected cell, instead of the normal yield of around 200, indicating that the 6-azauracil effectively inhibits further RNA synthesis in complete medium, although the RNA which was preformed in the adenine medium remains productive.

Discussion.—It is our hypothesis that a specific RNA accumulates under cultural conditions that inhibit protein synthesis, and that such RNA is required for ultimate intracellular phage production. The question arises whether some compound or property other than RNA is responsible for this acquired characteristic that permits T2 production under conditions that ordinarily will not support phage synthesis. Although none of the other possibilities has been completely excluded, examination of the data tends to negate their assignment as the causative factor in this phenomenon. The possibility that endogenous adenine is the responsible agent is largely eliminated by those experiments with prolonged intermediate incubation in minimal medium, or medium containing deoxyadenosine alone. The subsequent response in medium containing arginine:deoxyadenosine was not altered. Likewise, incubation of uninfected B94 in adenine alone, which may be expected to accumulate endogenous adenine to approximately the same extent as T2 infected bacteria, does not create conditions that will elicit phage synthesis in the deoxyadenosine:arginine medium. These same arguments may be used (although with less assurance) to rule out some free nucleotide, or other low molecular weight compound, as the responsible factor.

Another possible source of adenine, or adenine derivatives, could arise from the breakdown of accumulated RNA when the cells are placed in the medium containing arginine and deoxyadenosine. The formation of labile RNA has been described by Neidhardt and Gros[33] upon chloramphenicol inhibition of *E. coli* B. It will be recalled, however, that uninfected B94, cultured in medium containing adenine alone, does not have the capacity to allow subsequent phage formation in arginine:deoxyadenosine medium, in spite of the fact that RNA labeling in preincubated cells attains the same level as in their T2-infected counterparts (Table 5). It is also unlikely that the RNA accumulated in phage-infected B94 functions by serving as an obligatory precursor for T2 DNA, inasmuch as the maximum synthesis of RNA represents only five to six T2 nucleic acid equivalents, but subsequent T2 yields in deoxyadenosine:arginine are around 20 (Table 3 and Fig. 3).

Adenine might be obtained from deoxyadenosine as a result of the induction of a deoxynucleoside phosphorylase (or deoxynucleosidase) created by preincubation in adenine. In order for this to be compatible with the data, it must be assumed that such enzyme activity is induced by adenine preincubation of only the phage infected cells, since preincubated, uninfected B94 do not respond in deoxyadenosine:arginine. It is difficult to reconcile this mechanism with the observation th t T2 production reaches a limit after about 30 min incubation in deoxyadenosine:arginine, since the deoxyriboside is available in the medium in large excess (40 μg/ml). Experiments using labeled deoxyadenosine are now under way in our laboratory to test this possibility, however.

DNA synthesis, as measured by P^{32} incorporation, appears to be quite limited during the incubation in adenine (Table 2), and the manner in which P^{32} is incor-

porated into DNA does not seem to bear any relationship to the final phage yield (Fig. 3). No increase in hydroxymethylcytosine is observable during the preincubation of T2-infected cells in adenine alone.

A distinct characteristic of these experiments is the fact that after adenine preincubation, the yield of phage in deoxyadenosine:arginine medium is limited, reaching a maximum after about 30 to 35 min incubation in the latter medium. Whatever functional role is assigned to RNA in phage formation, therefore, must account for the fact that the activity of this RNA becomes rather quickly exhausted; with this organism, it can be restored by the addition of adenine (Fig. 1). These observations are not easily reconcilable with a mechanism of RNA function based on a *simple* template or carrier action. It is tempting to speculate that what we are observing is the formation of a specific RNA that is responsible for the formation of specific proteins-phage protein itself or newly induced enzymes.[34-36] It seems unlikely that the RNA would be required exclusively for the formation of enzyme protein, since such enzymes would not be expected to lose activity so abruptly 35 min after their formation. On the other hand, a certain amount of RNA may be associated with the synthesis of a limited quantity of phage protein, but during the synthesis of the protein the RNA is rendered nonfunctional. Such an interpretation would account for the limited phage production after removal of adenine, and upon presenting the metabolites required for protein and DNA syntheses. It may be noted that after the consecutive removal of adenine, then arginine, there appears to have been present a reserve of phage protein that could be later utilized upon addition of deoxyadenosine (Fig. 2, curve B).

Summary.—The *E. coli* mutant B94 requires adenine and arginine both for growth and ability to produce T2 bacteriophage. While the mutant does not synthesize phage in medium in which deoxyadenosine is substituted for adenine, it can do so if the infected cells are first incubated in medium containing adenine only. The ultimate yield of bacteriophage is reasonably correlated with the extent of synthesis of a specifically labeled RNA that accumulates during incubation of infected cells in the medium containing adenine alone. Uninfected B94, cultured in adenine medium, incorporates a similar level of radioactivity into the free nucleotide pool, DNA and RNA (labeled in a characteristically different manner from the RNA of comparably T2-infected B94): these cells, however, are unable to produce phage in medium containing deoxyadenosine:arginine. A tentative conclusion from these data is that the synthesis of a *specific* RNA is required for intracellular bacteriophage production, before phage protein and DNA synthesis can proceed.

* Operated by Union Carbide Corporation for the U. S. Atomic Energy Commission.

† A slant of B94 was generously supplied by Dr. Joseph Gots of the University of Pennsylvania School of Medicine.

[1] Putnam, F. W., in *Advances in Protein Chemistry* (New York: Academic Press, 1953), vol. 8, p. 177.

[2] Volkin, E., and L. Astrachan, *Virology*, **2**, 594 (1956).

[3] Cohen, S. S., *J. Biol. Chem.*, **174**, 281 (1948).

[4] Kozloff, L. M., K. Knowlton, F. W. Putnam, and E. A. Evans, Jr., *J. Biol. Chem.*, **188**, 101 (1951).

[5] Manson, L. A., *J. Bacteriol.*, **66**, 703 (1953).

[6] Volkin, E., and L. Astrachan, *Virology* **2**, 149 (1956).

[7] Hershey, A. D., *J. Gen. Physiol.*, **37**, 1 (1953).

[8] Volkin, E., in *Fourth International Congress of Biochemistry* (New York: Pergamon Press, 1959), vol. 7, p. 212.
[9] Volkin, E., L. Astrachan, and J. L. Countryman, *Virology*, **6**, 545 (1958).
[10] Astrachan, L., and E. Volkin, *Biochim. et Biophys. Acta*, **29**, 536 (1958).
[11] Countryman, J. L., and E. Volkin, *J. Bacteriol.*, **78**, 41 (1959).
[12] Volkin, E., and L. Astrachan, in *The Chemical Basis of Heredity*, eds. W. D. McElroy and B. Glass (Baltimore: The Johns Hopkins Press, 1957) p. 686.
[13] Nomura, M., B. D. Hall, and S. Spiegelman, *Fed. Proc.*, **19**, Part 1, 315 (1960).
[14] Watanabe, I., Y. Kiho, and K. Miura, *Nature*, **181**, 1127 (1958).
[15] Jeener, R., *Biochim. et Biophys. Acta*, **27**, 665 (1958).
[16] Pardee, A. B., and L. S. Prestidge, *Biochim. et Biophys. Acta*, **37**, 544 (1960).
[17] Astrachan, L., and E. Volkin, *Biochim. et Biophys. Acta*, **32**, 456 (1959).
[18] Volkin, E., unpublished observations.
[19] Gollub, E. G., and J. S. Gots, *J. Bacteriol.*, **78**, 320 (1959).
[20] Herriott, R. M., and J. L. Barlow, *J. Gen. Physiol.*, **36**, 17 (1952).
[21] Adams, M. H., in *Methods in Medical Research*, ed. J. H. Comroe, Jr. (Chicago: Year Book Publishers, 1950), vol. 2, pp. 1–73.
[22] Hershey, A. D., J. Dixon, and M. Chase, *J. Gen. Physiol.*, **36**, 777 (1953).
[23] Tyner, E. P., C. Heidelberger, and G. A. LePage, *Cancer Research* **13**, 186 (1953).
[24] Cohn, W. E., *J. Am. Chem. Soc.*, **72**, 1471 (1950); also *J. Cell. and Comp. Physiol.*, **38**, Suppl. 1, 21 (1951).
[25] Griswold, B. L., F. L. Humoller, and A. R. McIntyre, *Anal. Chem.*, **23**, 192 (1951).
[26] Astrachan, L , unpublished observations.
[27] Ernster, L., R. Zetterstrom, and O. Lindberg, *Acta Chem. Scand.*, **4**, 942 (1950).
[28] Evans, E. A., Jr., *Fed. Proc.*, **15**, 827 (1956).
[29] Colowick, S. P., and N. O. Kaplan, eds., *Methods in Enzymology* (New York: Academic Press, 1955), vol. 2, pp. 448, 456, and 468.
[30] Hershey, A. D., and N. E. Melechen, *Virology*, **3**, 207 (1957).
[31] Tomizawa, J., and S. Sunakawa, *J. Gen. Physiol.*, **39**, 553 (1956).
[32] Burton, K., *Biochem. J.*, **61**, 473 (1955).
[33] Neidhardt, F. C., and F. Gros, *Biochim. et Biophys. Acta*, **25**, 513 (1957).
[34] Kornberg, A., S. B. Zimmerman, S. R. Kornberg, and J. Josse, these PROCEEDINGS **45**, 772 (1959).
[35] Flaks, J. G., J. Lichtenstein, S. S., Cohen, *J. Biol. Chem.*, **234**, 1507 (1959).
[36] Bessman, M. J., *J. Biol. Chem.*, **234**, 2735 (1959).
[37] Sommerville, R. and G. R. Greenberg, *Fed. Proc.*, **13**, 327 (1959).
[38] Keck, K., H. R. Mahler, and D. Fraser, *Arch. Biochem. Biophys.*, **86**, 85 (1960).

---

*A COMPARISON OF ANIMAL HEMOGLOBINS BY TRYPTIC PEPTIDE PATTERN ANALYSIS**

BY EMILE ZUCKERKANDL,† RICHARD T. JONES, AND LINUS PAULING

DIVISION OF CHEMISTRY AND CHEMICAL ENGINEERING,‡ CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated August 22, 1960

The complete amino acid sequences (primary structure) of hemoglobins can, in principle, be determined by methods currently available. Although detailed studies of the primary structure of human and horse hemoglobins are in progress in several laboratories,[1] the methods are so laborious that complete sequences have not yet been established. Important questions in the realm of genetics and evolution require the immediate examination of the structure, primary and other, of

many different hemoglobins. The application of methods that are quicker, though less informative and reliable, than the techniques required for complete sequence determination is therefore in order as a provisional means of securing useful information. Such a method is the analysis of peptide patterns obtained by combined paper electrophoresis and chromatography of tryptic hydrolysates of denatured hemoglobin.[2] Of particular interest are comparisons between hemoglobin components present in (*a*) organisms of one animal species at a given time in development, (*b*) organisms of one species at different stages of development, and (*c*) organisms of different species. The present paper is concerned exclusively with the last type of comparison. In order to scan the range of variation of hemoglobin structure throughout evolution, hemoglobins from a number of animals both closely and distantly related to man have been selected and compared as to tryptic peptide patterns with human hemoglobin A. Whole hemoglobin preparations from adult animals have been studied throughout. The problem of individual heterogeneity will be treated elsewhere.

Materials and Methods.—Ape bloods[3a] were anticoagulated and transported in Alsever's solution. The apes studied included two lowland gorillas (*Gorilla gorilla*), two chimpanzees (*Pan troglodytes*), and three orangutans (*Pongo pygmaeus*). Erythrocytes from Rhesus monkeys (*Macaca mulatta*) were obtained from clotted blood.[3b] Heparinized porcine and bovine bloods were secured during bleedings at Los Angeles slaughter houses. Heparinized bloods were obtained from the marine lungfish, *Pimelometopon pulcher* (sheepshead), and from the cartilaginous fish *Cephaloscyllium uter* (swell shark).[3c] Blood was also obtained from a live specimen of *Lepidosiren paradoxa* (Dipneust, South American fresh water lungfish).[3c, d] Contamination of the latter blood by tissue fluid was unavoidable. Blood of the Pacific hagfish, *Polistotrema stouti*, was also examined.[4] Coelomic cells (hemoglobin cells) were obtained from *Urechis caupo* (Echiurid marine "worm")[5] collected at low tide from mud flats near the Kerckhoff Marine Laboratory at Corona del Mar, California.[3e]

In general, the hemoglobin preparations examined were from single individuals. However, in the case of *Urechis* single as well as pooled samples of coelomic cells were used without apparent differences. Only a pooled sample of blood from twelve hagfish was examined. All of the animals studied were judged to be adults either from their size or their known age. The youngest apes were one orangutan 2 years old and one gorilla 2 years old. Peptide patterns of their hemoglobins were indistinguishable from those of older individuals of the same species. The *Lepidosiren*, 14 inches long, was at a minimum 8 months of age, and might be considered a subadult.

The erythrocytes were washed four times with cold 0.9% NaCl with the exception of *Pimelometopon* and *Cephaloscyllium* cells, which were washed with 1.2% NaCl.[6] The red cells from hagfish were washed with chilled 3% NaCl and the coelomic cells of *Urechis* with 3.2% NaCl. Washing with 0.9% NaCl led to considerable hemolysis in the case of the fresh water fish *Dipneust*.

The washed cells were hemolyzed in a standard fashion with distilled water and toluene[7] with the exception of *Urechis*, where ether was used instead of toluene. The washed red cells of hagfish were stored for one day at 4°C before lysing. A major portion of the hemoglobin obtained was insoluble possibly due to acidifica-

tion during storage. The peptide pattern shown below was obtained from the remaining soluble fraction. After lysis, each sample was centrifuged at high speeds in order to remove solid debris and the toluene phase. The resulting hemoglobin solutions were saturated with carbon monoxide and dialyzed for a week against distilled water saturated with carbon monoxide (three changes of water with the ratio of hemoglobin solution to water of the order of 1:100). This prolonged dialysis was employed to permit the flocculation of non-hemoglobin proteins.[6] A crystal of thymol was added to one *Urechis* preparation in order to eliminate the possibility of the formation of peptides by bacterial contamination. No differences were found in the peptide patterns obtained with and without thymol.

With species closely related to man the hemoglobin concentration in the final preparation could be assayed approximately by the use of spectrophotometric constants established for human hemoglobins. In other species the quantities of material to be used on peptide patterns were estimated roughly from the consumption of sodium hydroxide during tryptic digestion.

Hydrolysis with trypsin of heat-denatured hemoglobin preparations was performed at constant pH using a Radiometer automatic titrator as a pH-stat. The hydrolyses were carried out at 40°C at pH 8.0 in the presence of Ca ion (0.01 *M*) with an enzyme to hemoglobin ratio of between 1 and 2% by weight (Worthington twice crystallized trypsin). Ninety minutes was allowed for the hydrolysis except in the case of the apes, where the digestion was stopped after sixty minutes.

Those tryptic peptides which are soluble at pH 6.5 were analyzed by a combination of electrophoresis and chromatography on paper. These peptide patterns were obtained essentially as specified by Ingram[2] except that the voltage rate was 900–1,000 V for a duration of $3^1/_2$ hours. In general, a human control sample was run as a pair mate with every animal hemoglobin in order to have a reference pattern for each unknown. Even with such parallel determinations, the spreading occurring during electrophoresis was not always identical. The polarity of the electrophoretic field is marked on each pattern and the point of application of the sample is designated by a cross (+). The peptide spots resulting from tryptic hydrolysis of human hemoglobin have been assigned arbitrary numbers by Ingram.[2] This convention has been employed in the present study. The neutral band region is comprised of peptide spots 1 through 7. These peptides do not migrate to any extent in the electric field, but do move some from the point of application due to fluid movement. Phenylalanine and lysine have been added as reference spots at least on one pattern in the case of every species. Chromatography was ascending in direction and the peptides were detected mainly by their reaction with ninhydrin. The Sakaguchi[8] test for arginine and the cinnamaldehyde[9] reaction for trypophan were also employed. Patterns from each species were obtained at least in triplicate.

Results.—Figure 1 is a peptide pattern of human hemoglobin A as currently obtained in our laboratories. A diagram indicating the peptide spot numbering system is also included.

Figure 2 shows peptide patterns from three different apes (gorilla, chimpanzee, and orangutan) and one monkey (Rhesus). The gorilla, chimpanzee, and human patterns are almost identical in appearance. In the case of the gorilla peptide patterns, spot numbers 12 and 24 both appear to be double. In humans the peptides that make up these two spots are from the beta chain. No difference was

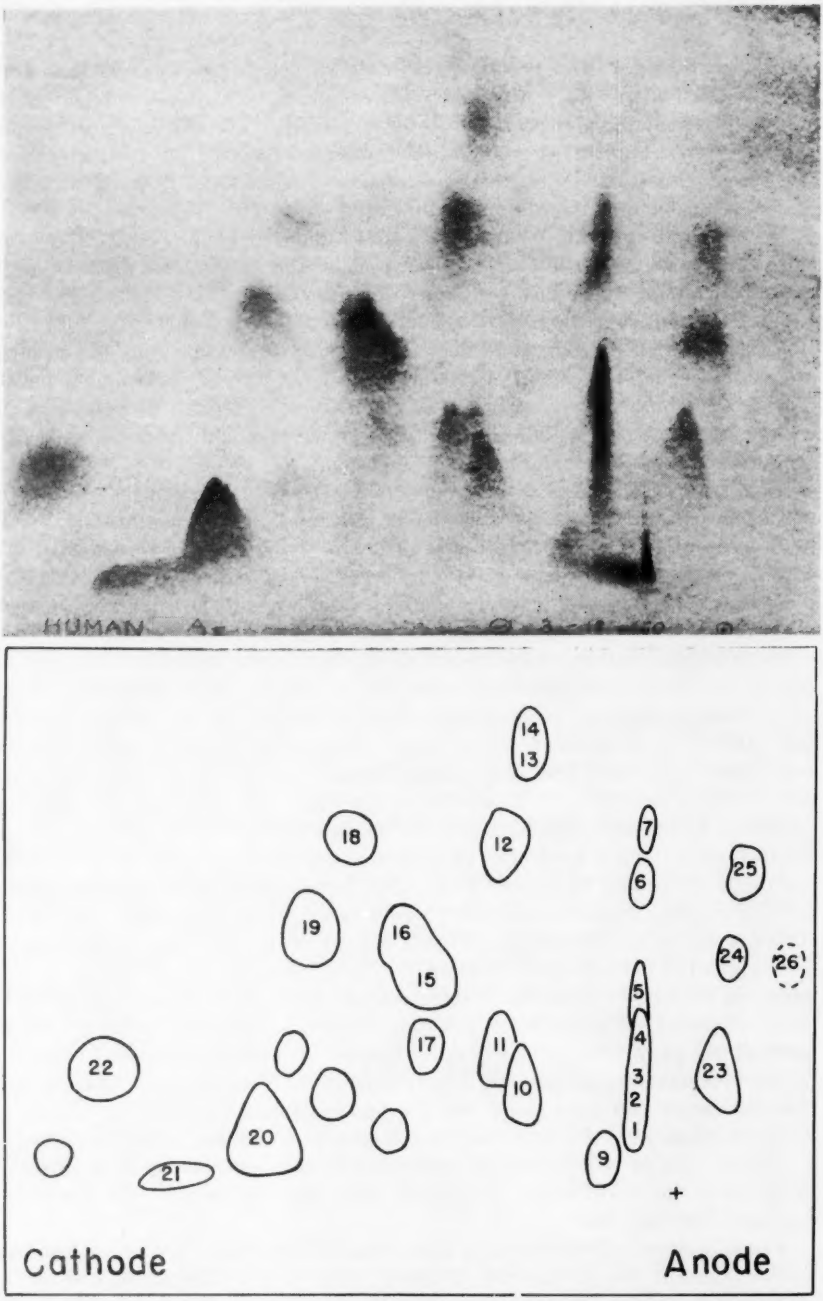


FIG. 1.—A tryptic peptide pattern of adult human hemoglobin and its schematic representation. The spots are numbered according to the convention introduced by Ingram.[2] The site of spot 26 is shown on the diagram, although not seen on the pattern represented. The locations of four hitherto undescribed spots usually found on patterns obtained in our laboratories are indicated.

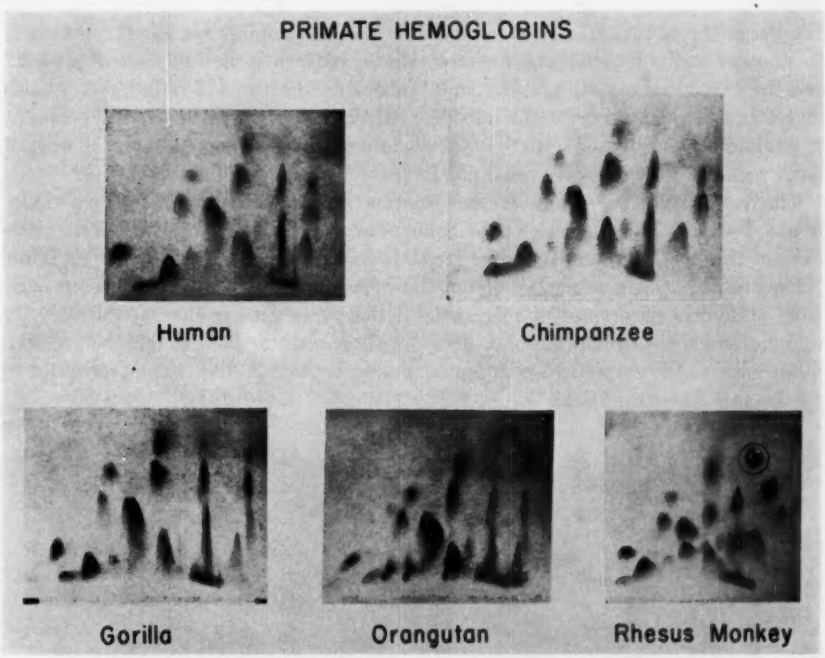


FIG. 2.—Tryptic peptide patterns of primate hemoglobins. The circled spot on the Rhesus monkey pattern represents phenylalanine added two and a half inches to the anodal side of the point of application of the peptide mixture.

observed between two gorillas, one male and one female. Spots 12 and 24 also seem to be double in the peptide patterns of chimpanzee. In addition, extra material may be present in the 15–16 region. In humans at least four major peptides are present in this region.[1] Two chimpanzees, both females, were found to have identical patterns. Of course further differences between these two types of apes and humans may be discovered upon the analysis of the individual peptides or the study of the protein residue which remains insoluble after hydrolysis with trypsin.

The difference from human patterns is somewhat greater for orangutan peptide patterns than for the patterns of the two apes just mentioned. In two orangutans (one male and one female) spots 12 and 24 are double. The third orangutan (female), whose pattern is illustrated in Figure 2, was observed to have single 12 and 24 spots. Further differences between the patterns of the individual orangutans are not apparent; however, each differs from human patterns by an apparent increase in peptides in the 15–16 region. The orangutans also differ from humans and the other two apes studied by the appearance of two new spots, one anodal to spot 20 and the other cathodal and below spot 10. The latter does not appear to be the same as spot 9 of human. All of the spots containing arginine and tryptophan are the same as in human.

Spot 12 is absent in the peptide patterns of Rhesus monkey hemoglobin. A new spot, possibly representing a modification of the peptide of spot 12, is present

just below the 12 region. The new spot, like 12 from human, contains tryptophan. In another individual Rhesus monkey, whose pattern is not presented, spot 12 and the new spot are both present simultaneously. Patterns of both monkeys also reveal another spot on the cathodal side of the 2, 3, and 4 region. Spot 7 is absent or modified (almost fused with 6). Other differences may be present in the neutral band region, which is poorly resolved.

The human spot 26, which contains arginine and is derived from the beta chain, is not always seen because of poor color production with ninhydrin. The existence of this spot can, however, be ascertained by the Sakaguchi test for arginine. The primate patterns presented do not show spot 26, either because it did not react with ninhydrin or because it migrated off the paper during electrophoresis. By undimensional electrophoresis and the application of the Sakaguchi test the presence of arginine in the zone corresponding to human peptide 26 has been ascertained in all the primates examined.

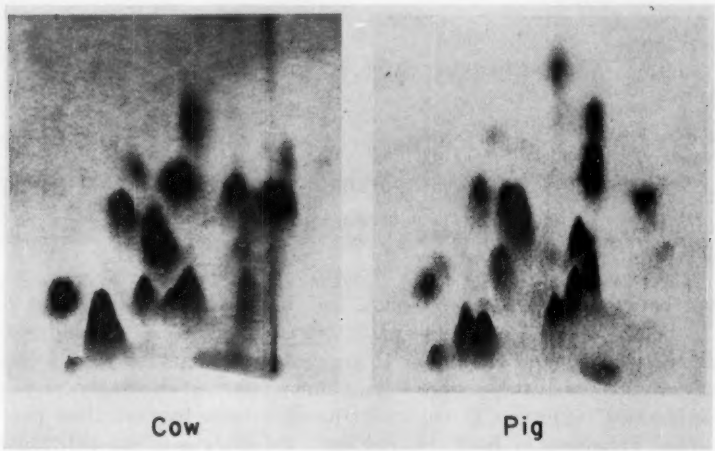


FIG. 3.—Tryptic peptide patterns of bovine and porcine hemoglobins.

Tryptic peptide patterns of bovine and porcine hemoglobins are presented in Figure 3. Cow and pig belong to the same order. In many respects these patterns appear similar to human patterns. However, the differences are numerous and striking.

In bovine and porcine hemoglobin, sequence val-leu has been found to be N-terminal in one of the polypeptide chains.[10] Both hemoglobins therefore contain alpha chains, according to the terminology introduced by Rhinesmith, Schroeder, and Martin.[11] However, the cow alpha chain is different in several respects from the human alpha chain. Although the chains have not yet been separated, the peptide patterns indicate that the alpha chain peptides 5, 10, and 17 are absent from bovine patterns, while the alpha chain peptides 12, 13, 18 and an alpha component of spot 16, detected by a positive Sakaguchi test, are present.[1,12] Although the second chain by definition is not a beta chain,[10, 11] the human beta chain spots

12 and 19 are seen clearly. The human pure beta spots 2, 4, 14 (as ascertained by a negative reaction with cinnamaldehyde) and 25 are absent.

The porcine alpha chain may be somewhat more similar to the human alpha chain since the alpha spots 10, 13, and 18 are seen, and the presence of the alpha spots 11, 17, and 23 is suspected. The presence of an alpha-component in the composite spot 15 is likely because of a positive cinnamaldehyde test, and the presence of an alpha-component in the composite spot 16 is indicated by a positive Sakaguchi test. The alpha spot 5 is absent. Among the beta spots, 14 (cinnamaldehyde positive), 26 (Sakaguchi positive), 4, 19, 24, and possibly 25 are present; while 12 is absent. Like bovine hemoglobin, porcine hemoglobin does not contain beta chains; however, there are a number of tryptic peptides similar to those from the human beta chain observed.

The composite human spots 13–14, 15–16, 20 and 21 as well as the lysine spot 22 (arising from both α and β chains[1]) are seen in pig as well as in cow. Thus the

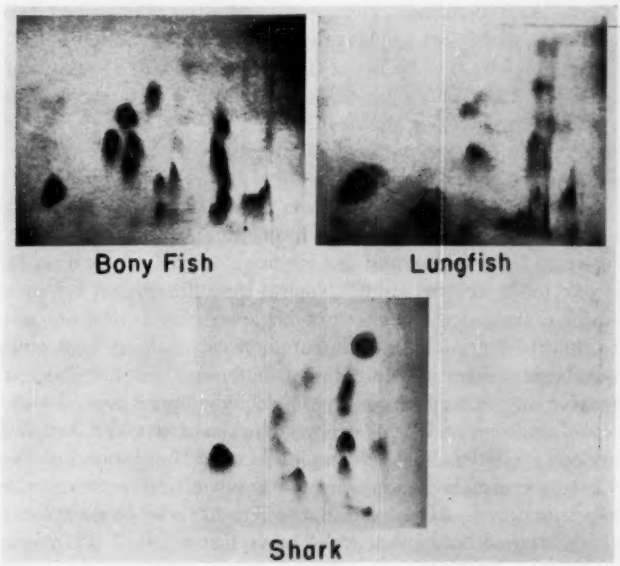


FIG. 4.—Tryptic peptide patterns of "fish" hemoglobins. The circled spot on the shark pattern represents phenylalanine added two and a half inches to the anodal side of the point of application of the peptide mixture.

apparent similarities between bovine and human patterns are in many respects the same as those noted between porcine and human; however, the differences are not the same. From three independent pattern studies of the same bovine sample it appears that many of the new spots are slightly on the anodal side of the neutral band region, whereas in porcine patterns most of the new spots are on the cathodal side. Arginine is present in at least two of the new porcine peptides, one neutral and one acidic.

Spot 12, as stated, and as confirmed by cinnamaldehyde tests, is absent from pig, but present in the cow. The decapeptide[1] that accounts for this spot in humans may have undergone independent alteration (mutation?) at different times. In pig, a novel tryptophan peptide is found in the upper neutral region. It may correspond to an altered peptide No. 12.

The three fish from which peptide patterns have been obtained (Fig. 4) belong to entirely different groups which modern systematicians do not unite under the heading of a single class. None of the three patterns is similar to the patterns of human hemoglobin. In the bony fish (*Pimelometopon*), spot 22 is present as are spots 18 and 19, which in humans are dipeptides.[1] Possibly spots 24 and 25 and some material in the regions of 4 and 15 are common to human and *Pimelometopon*. The remaining spots appear to be different from human. Three spots containing tryptophan are seen with no analogy in human hemoglobin. One other spot at the base of the neutral band resembles a similarly located spot on pig patterns in that it gives a blue rather than purple color with ninhydrin and a positive Sakaguchi test. In the shark (*Cephaloscyllium*), spot 22 is again present as are probably spots 4, 5, and 7. A spot is present in the region of peptide 12, but differs from spot 12 as shown by the absence of tryptophan. Spots 18 and 19 are absent as well as spots 20 and 21. Four peptides containing tryptophan are seen with no analogy in human or in *Pimelometopon*.

In the case of the peptide patterns of lungfish (*Lepidosiren*) hemoglobin, spot 22, (lysine, identified by position and shade of color) is present. Spots 19 and 23 may also be present. The absence of spots 20 and 21 appears to be characteristic of all three types of fish. Three Sakaguchi-positive peptides are seen in the middle and upper neutral band region and one cinnamaldehyde positive peptide adjacent to the lower part of the neutral band. None of these are present in human. Among the vertebrates examined, *Lepidosiren* hemoglobin may be the one with the least number of primary structural traits in common with human hemoglobin.

There may be some peptides in common between *Cephaloscyllium* and *Pimelometopon*, notably one Sakaguchi-positive spot in the upper neutral band and some between *Cephaloscyllium* and *Lepidosiren*. *Pimelometopon* and *Lepidosiren* appear to have only one ninhydrin spot in common beyond the lysine spot, No. 2[1].

The Cyclostomes are the most primitive group of living vertebrates, although they are very specialized. The molecular weight of those Cyclostome hemoglobins that have been studied is reported to be one quarter[13, 14] of the molecular weight of other vertebrate hemoglobins. Therefore it might be presumed that these Cyclostome hemoglobins are composed of single polypeptide chains. Because direct estimates of the molecular weight of *Polistotrema* hemoglobin are not available, an ultracentrifuge study was made.[15] The soluble fraction of our *Polistotrema* hemoglobin (see *Materials and Methods*) appeared to sediment as a single component with an $s_{20,w}$ of 1.9. This value, in the case of hemoglobin from the related *Petromyzon*, has been interpreted by others[14] to correspond to a molecular weight of approximately 23.600. However, the peptide pattern of *Polistotrema* hemoglobin shown in Figure 5 may not represent a single kind of polypeptide chain because two major and possibly some minor hemoglobin components were shown to be present by starch gel electrophoresis.[16]

It is not possible to make definite statements about the similarities between

Polistotrema and other patterns in the region of the neutral band because of incomplete resolution of the *Polistotrema* spots. Except possibly for the neutral band region and the lysine spot, the *Polistotrema* peptide pattern has no feature in common with human, lungfish, or shark patterns. One or two spots in the bony fish pattern appear similar to spots in the *Polistotrema* pattern.

In order to test the range of variation in hemoglobins in general, a peptide pattern of a single invertebrate, the Echiurid "worm" *Urechis* was studied (Fig. 6). In this group the hemoglobins are intracellular, as in the vertebrates. The heavily

FIG. 5.—Tryptic peptide pattern of hagfish hemoglobin. The circled spots represent phenylalanine and lysine respectively added at the point of application of the peptide mixture.

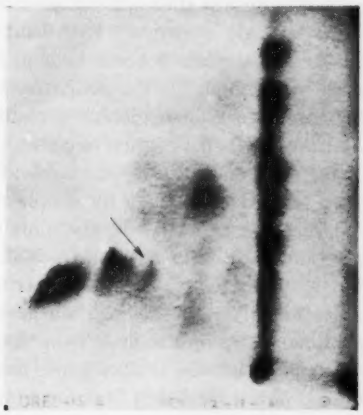


FIG. 6.—Tryptic peptide pattern of hemoglobin from an Echiurid "worm." Most of the spots in the neutral band, as well as the one indicated by the arrow, are not tryptic peptides (see text).

loaded neutral band region obtained has been shown to contain many amino acids[17] that are not the result of tryptic hydrolysis. The spots on the cathodal side of the neutral band, with the exception of the one indicated by the arrow, were found to be from *Urechis* hemoglobin. The latter conclusion was obtained by comparing, peptide patterns of the total undigested hemoglobin preparation with the undigested supernatant obtained after heat denaturation and centrifugation of the precipitate. The only spot identified with any certainty is spot 22. All other *Urechis* spots appeared to be without a match in human.

Discussion.—Peptide patterns as obtained by the present technique do not yield unequivocal information for several reasons: (*a*) spots found in identical locations may not represent identical peptides, (*b*) single spots may contain more than one peptide, (*c*) a single peptide, through partial secondary alteration or because of incomplete splitting into two peptides, may give rise to more than one spot, and (*d*) a significant part of the hemoglobin molecule (about one third in human A, S, and F[18]) is not broken down by trypsin sufficiently to become soluble, and accordingly is not represented on tryptic peptide patterns.

Although, in view of (*d*), all findings must be substantiated by the examination of the undigested core, observations (*a*), (*b*), and (*c*) do not seriously impair the conclusion that Primate and human hemoglobins are very similar, especially

gorilla, chimpanzee and human. There is no doubt that the identification of individual spots needs to be based on data more specific than position, shape, and (sometimes) color shade of the ninhydrin-developed spots, or even reaction with reagents specific for certain amino acid residues. However, when two complex peptide patterns look similar as a whole the probability that most of the spots actually represent identical or highly similar sequences becomes high.

That chimpanzee should be extremely close to man while orangutan and Rhesus monkey are somewhat more distant when examined by hemoglobin peptide patterns is in accord with serological data.[19] On the other hand, while human hemoglobin strongly crossreacts with Rhesus monkey hemoglobin, it has been observed not to do so with porcine hemoglobin and only exceptionally crossreacts with cattle hemoglobin.[20] The similarity between human and Primate hemoglobins suggests that these hemoglobins probably have not been modified extensively since the times of their common ancestor. It is possible that the genes for the α and β chains of a normal adult human hemoglobin are more stable than the mutated genes formed from them by simple mutations, and that the mutated genes often undergo back-mutation to the more stable normal genes. In addition to natural selection, the thermodynamic stability of the genes themselves may be an important factor in determining the distribution of alleles in a population. Our results indicate that stable hemoglobin genes had been developed before the separation of humans and anthropoids from their common stock.

The idea that the normal genes have greater thermodynamic stability than their mutant alleles, and that in consequence the mutant alleles have a greater mutation rate than the normal genes, provides an explanation of the fact that some of the abnormal hemoglobins have their abnormalities at the same site.[21] The alleles corresponding to one site may represent one mutant from the normal gene and other alleles formed by second mutations from this mutant, and involving its unstable site.

Human adult hemoglobin peptide patterns actually differ more from human fetal patterns than from adult gorilla and chimpanzee patterns. In view of the theory of recapitulation it would not appear paradoxical that an adult organism should be in a sense a more distant "relative" of its own embryo than of other closely related adult organisms. Adaptative factors in response to the environment may play an equally important role in establishing the differences. In the only animal examined (sheep) fetal hemoglobin differs from adult hemoglobin in molecular shape.[22]

As one gets further away from the group of Primates, the amount of primary structure that is shared with human hemoglobin decreases. Different primary structures may be compatible with rather constant tertiary structures.[23] The limits to this statement remain to be investigated.

The observation that pig and human on one hand and steer and human on the other show to a large extent apparent identity of the same peptides but diverging differences for the others may again indicate the existence in the hemoglobin molecule of zones more prone to mutation than others, or of zones where the occurring mutations are more liable to be preserved by natural selection. Part of the Primate pattern may have evolved relatively early in mammalian history, if not before, while another part may have varied more frequently throughout the groups.

Sections of both the alpha and beta chains, or of chains taking their place, appear to be involved. So far there is no evidence that one of the chains remains more stable through evolution than the other.

Firm conclusions about the partial stability of chains are premature before the sequence of each of the tryptic peptides that are common to different hemoglobins is determined. On the basis of a chance distribution of amino acids in the hemoglobin molecule, it is evident that the smaller the peptide, the greater the chance of finding it in different species. The presence of spot 22 seems to be the only trait shared by all hemoglobins. Because 22 is pure lysine, this only means that in all of the hemoglobins examined there is, in at least one place, a lysine next to another lysine or an arginine. In human hemoglobin a lys-lys sequence has been found to occur in both the alpha and beta chains.[1] The frequently seen spots 18 and 19 are dipeptides.[1] Spots 20 and 21 are a tetrapeptide and a pentapeptide, respectively, which are closely related.[1] It is true that some of the spots that may be common to pig, steer, and human (spots 10, 11, 24, and 25) contain much higher peptides.

However, the significance of the similarities observed between pig, steer, and human hemoglobins is increased by the observation that many of these similarities are not shared by the three fish and none of them (except the lysine spot) by the Cyclostome and the *Urechis* "worm." The three "fish" patterns differ among themselves considerably more than the mammalian patterns examined. Although shark and bony fish as well as shark and lungfish hemoglobins may share a small number of tryptic peptides, bony fish patterns appear about as dissimilar from lungfish as from human patterns. The findings are consistent with the view that these "fish" belong to widely divergent evolutionary lines.

Unless a greater constancy were found in the undigested core than in the rest of the molecule, no large tryptic peptide would be constant throughout the vertebrate series. Thus, most of the hemoglobin molecule, at least in the portion solubilized by trypsin, has been subject to successful mutation at some time during vertebrate evolution. Successful mutations in each tryptic peptide region have probably occurred more than once considering the differences between groups of fish and between Cyclostomes and fish. This observation, to be sure, does not exclude the existence of preferential sites of (successful) mutation. But such sites, if they exist, must be distributed throughout the molecule and not be concentrated in any one part of it, in contradistinction to what has so far been observed with insulin.[24]

* This work was supported in part by Grant No. H3136 from the National Institutes of Health, U.S. Public Health Service, and was presented in part at the 138th meeting of the American Chemical Society.

† On leave from Centre National de la Recherche Scientifique, Paris.

‡ Contribution No. 2618.

[1] Groups investigating the amino acid sequence of human hemoglobin include Dr. G. Braunitzer and co-workers, Drs. R. J. Hill and W. Koingsberg, Drs. V. M. Ingram and J. A. Hunt, and Dr. W. A. Schroeder and co-workers. Information concerning the tryptic peptides is from unpublished sequence studies of W. A. Schroeder and co-workers.

[2] Ingram, V. M., *Biochim. Biophys. Acta*, **28**, 539–545 (1958).

[3] Various hemoglobin specimens have been supplied for this research. The authors wish to acknowledge the very valuable help in this respect of (*a*) Dr. W. P. Henschele, Head Veterinarian, San Diego Zoo; (*b*) Dr Renato Dulbecco, Division of Biology, California Institute of Tech-

nology; (*c*) Dr. K. S. Norris, Curator of Marineland of the Pacific, Palos Verdes, California; (*d*) Dr. W. H. Hildemann, Department of Zoology, University of California at Los Angeles, who drew the lungfish blood; and (*e*) Dr. A. Tyler, Division of Biology, California Institute of Technology.

[4] Specimens of the Pacific hagfish were caught and bled by Dr. David Jensen of the Oceanography Institute, La Jolla, California. The special methods employed were devised by him.

[5] Redfield, A. C., and M. Florkin, *Biol. Bull.*, **61**, 185–210 (1931).

[6] Roche, J., Y. Derrien, and M. S. Chouaiech, *Ann. Inst. Oceanograph.*, **20**, 97–113 (1940).

[7] Drabkin, D. L., *Arch Biochem.*, **21**, 224 (1949).

[8] Acher, R., and C. Crocker, *Biochim. Biophys. Acta*, **9**, 704 (1952).

[9] Method of J. Harley-Mason, *Biochem. J.*, **69**, 60 P (1958), modified by Kenneth N. F. Shaw (personal communication).

[10] Ozawa, H., and K. Satake, *J. Biochem.*, **42**, 641–647 (1955).

[11] Rhinesmith, Herbert S., W. A. Schroeder, and Nancy Martin, *J. Am. Chem. Soc.*, **80**, 3358 (1958).

[12] Hunt, J. A., *Nature* **183**, 1373–1375 (1959).

[13] Svedberg, The, *J. Biol. Chem.*, **103**, 311–325 (1933).

[14] Lenhert, P. G., W. E. Love, and F. D. Carlson, *Biol. Bull.*, **111**, 293–294 (1956).

[15] Made under the direction of Mr. John E. Hearst in the laboratory of Dr. J. Vinograd.

[16] Performed by Mr. Donald Sheffler in the laboratory of Dr. Ray Owen.

[17] The identification of these amino acids was made by Dr. Thomas L. Perry and Dr. Kenneth Shaw.

[18] Barrett, H. W., and W. A. Schroeder, unpublished.

[19] Mollison, Th., quoted by P. Kramp in *Primatologia* I, ed. H. Hofer, A. Schultz, and D. Starck (New York: S. Karker, Basel, 1956), pp. 1015–1034.

[20] Hektoen, L., and A. K. Boor, *J. Infect. Dis.*, **49**, 29–36 (1931).

[21] Ingram, V. M., *Brit. Med. Bull.*, **15**, 27–32 (1959), and unpublished work from this laboratory.

[22] Bragg, W. L., and M. F. Perutz, *Acta Cristall.*, **5**, 323–328 (1952).

[23] Perutz, M. F., M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, and A. C. T. North, *Nature*, **185**, 416–422 (1960).

[24] Harris, J. I., F. Sanger, and M. A. Naughton, *Arch. Biochem. Biophys.*, **65**, 427–438 (1956).

---

*THE KINETICS OF DOUBLE HELIX FORMATION FROM POLYRIBOADENYLIC ACID AND POLYRIBOURIDYLIC ACID**

BY PHILIP D. ROSS† AND JULIAN M. STURTEVANT

STERLING CHEMISTRY LABORATORY, YALE UNIVERSITY

Communicated by Raymond M. Fuoss, August 17, 1960

Synthetic polyadenylic acid (poly A) and polyuridylic acid (poly U), which are largely in random coil conformation in solution at neutral pH, have been shown to react to form a double-stranded helical complex, poly (A + U), which has a structure similar to that of naturally occurring deoxyribonucleic acid (DNA).[1–4] The reaction is accompanied by a decrease in the optical density at 259 mμ. Under most circumstances the reaction is too rapid for kinetic study by conventional spectrophotometric means, but it can be readily followed by the stopped-flow technique.[5, 6] We report here the results of measurements made by this technique at various polymer concentrations, ionic strengths, and temperatures.

The polymer samples employed were kindly supplied by Dr. A. Rich, who determined the mean sedimentation coefficients to be 6.25 for the poly A and 4.52

for the poly U. These values correspond to molecular weights of about 225,000 and 115,000 for poly A in the random coil form.[4] The polymer solutions contained 0.01*M* sodium cacodylate and were at pH 7.2, well above the pH at which poly A undergoes a conformational transition.[7] Concentrations, expressed as monomer molarities, were estimated by measurements of the absorbancy of the solutions both before and after alkaline hydrolysis.[3, 8] In mixing experiments at constant total concentration minimum absorbancy was observed at the expected equivalence point for the 1:1 complex. The temperature was controlled and measured with an accuracy of ±0.05°.

Results.—The change of optical density during the reaction does not follow

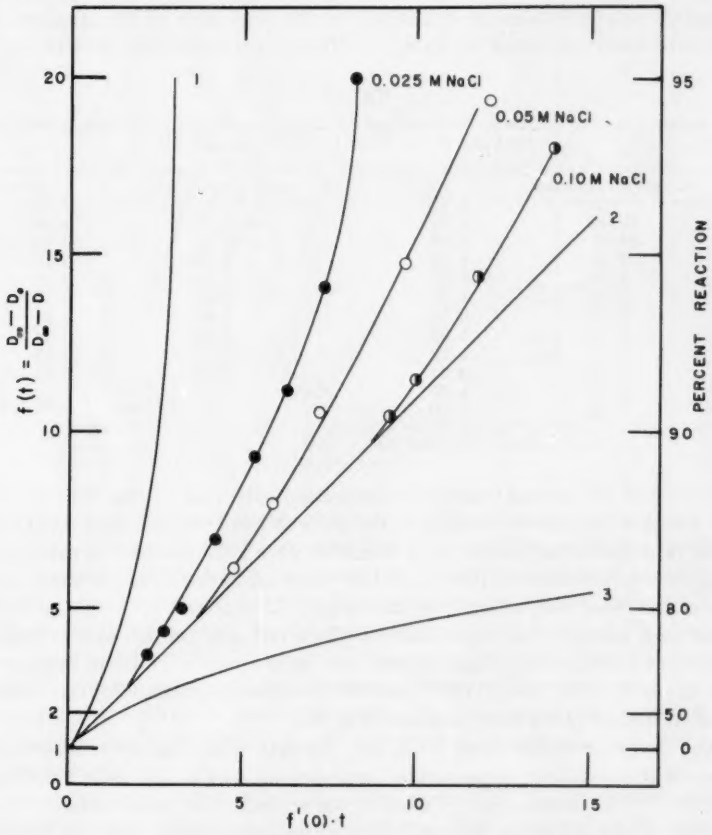


FIG. 1.—Dimensionless plot illustrating the course of the formation of the poly (A + U) complex at 25°. The reciprocal of the fraction not reacted is plotted against the product of the initial rate of change of this function by the time. The curves labeled 1, 2, and 3 are respectively the universal curves for simple first, second, and third order reactions (equal initial concentrations). Experimental observations at values of the abscissa less than 0.1 were generally made but are omitted from the graph for clarity. pH 7.2, 0.01 *M* sodium cacodylate, added salt as indicated; polymer concentration 10^{-4} monomoles per liter.

any simple kinetic expression. The data for typical experiments with equimolar polymer solutions at three different ionic strengths are presented in Figure 1, in the form of a dimensionless plot in which all irreversible reactions (and reversible reactions in the case of first order) of a given order follow a universal curve provided the initial reactant concentrations are equal. In this plot the ordinate is the reciprocal of the fraction unreacted and the abscissa is the product of the initial rate of change of the ordinate function by the time. At all salt concentrations, the decrease of optical density initially follows apparent second order kinetics, and then deviates toward first order behavior in the later stages. Similar curves were obtained at other polymer concentrations and temperatures. The rate during the second order part of the reaction is proportional to the product of the initial polynucleotide concentrations, as evidenced by the constancy of the apparent second order rate constants listed in Table 1. There is considerable uncertainty in the

TABLE 1

APPARENT SECOND ORDER RATE CONSTANTS FOR THE REACTION OF EQUIMOLAR SOLUTIONS OF POLY A AND POLY U AT 25° AND pH 7.2

0.10 *M* NaCl		0.05 *M* NaCl	
Polymer concentration monomer moles per l. × 10^5	10^{-5} *k*, l. mole^{-1} sec^{-1}	Polymer concentration monomer moles per l. × 10^5	10^{-5} *k*, l. mole^{-1} sec^{-1}
11.40	0.92	11.71	0.16
10.67	0.88	10.82	0.16
9.15	1.03	9.36	0.16
5.75	0.94	5.94	0.16
5.57	1.06	5.49	0.17
5.24	0.95	3.68	0.14
3.57	1.02	2.60	0.14
2.77	1.10	2.22	0.15
2.44	0.91	1.36	0.13
2.26	1.00		
1.25	1.00	Mean	0.15 ± 0.01
Mean	0.98 ± 0.06		

evaluation of the second order rate constants in the case of reactions at very low ionic strength because of bending of the plots at relatively low extents of reaction. Initial rates following second order behavior were also observed in experiments in which the concentration of either polymer was maintained constant while the molar ratio of the other was varied over the range 0.33 to 2.00.

The ionic strength has a pronounced effect, not only on the kinetic form of the reaction as illustrated in Figure 1, but also on the rate. Thus, an increase of the ionic strength at 25° from 0.015 *M* to 0.11 *M* causes a thousandfold increase in the apparent second order rate constant (Fig. 2).

Over the temperature range 5° to 55°, the rate of the change of absorbancy *decreases* with increasing temperature, approaching zero at the helix "melting temperature."[9] As illustrated in Figure 3, the second order rate constant is a linear function of the difference between the melting temperature and the temperature of the experiment, with the temperature coefficient dependent on ionic strength. Since the melting temperature increases with ionic strength (see Fig. 2), a small part of the variation of rate at constant temperature with ionic strength commented on above is to be attributed to this cause.

Discussion.—Although no direct evidence is available that the change in optical density is proportional to the extent of reaction during the formation of poly

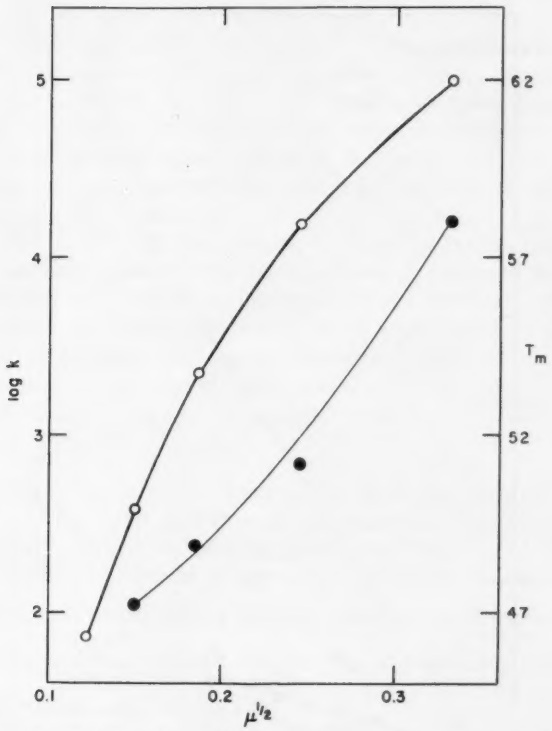


FIG. 2.—The effect of ionic strength on the apparent second order rate constant at 25° (open circles, liter mole^{-1} sec^{-1}) and the helix melting temperature (filled circles °C).

(A + U),[10] we shall assume that complications due to any lack of proportionality are negligible. The kinetic results reported here are consistent with the hypothesis that the formation of poly (A + U) involves the interaction of pairs of polynucleotide chains to form nuclei of a few hydrogen-bonded base pairs, followed by a cooperative growth of ordered helical regions from these nuclei. The marked effect of ionic strength on the initial rate of the reaction is in the direction expected for a reaction between polyelectrolytes carrying charges of the same sign. At high ionic strength the nucleation process remains rate limiting during most of the reaction, as indicated by the continuation of second order behavior until the reaction is nearly complete.

The complexity of the kinetics is particularly evident in the experiments at low ionic strengths, where deviations toward apparent first order behavior occur in the early stages of reaction. Three possible causes of this complex kinetic behavior can be suggested: (1) The rate of the cooperative ordering of the chains is more retarded by decrease in the ionic strength than the rate of nucleation, so that rate limitation is not confined to the nucleation process at low ionic strength. (2) Essentially all the bases in the polymer present in smaller concentration are paired at equilibrium,[12] so that realignment of mated chains must take place to eliminate

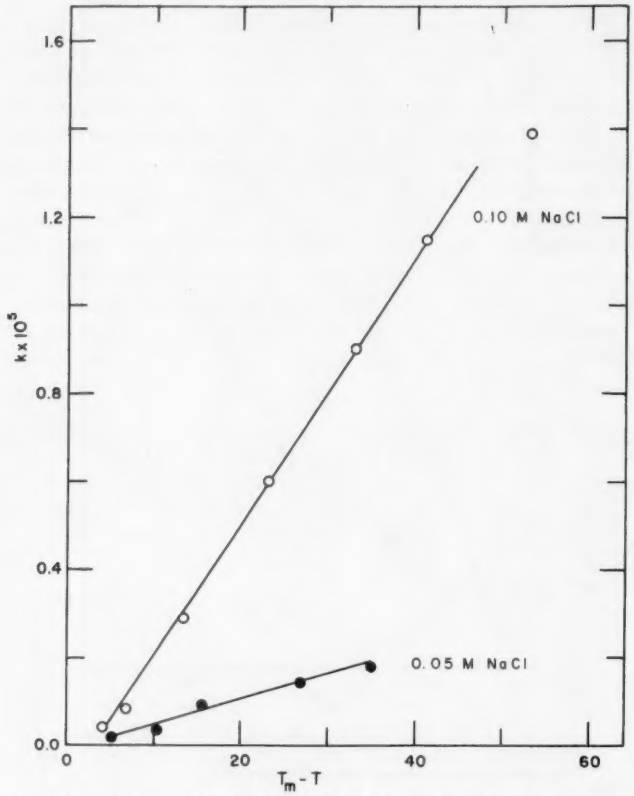


FIG. 3.—The variation of the apparent second order rate constant (liter mole^{-1} sec^{-1}) with the difference between the helix melting temperature and the temperature of the experiment, at two different ionic strengths. Polymer concentration 8.7×10^{-5} M.

unpaired regions left as a result of random initial formation of nuclei. This process would not be expected to follow apparent second order behavior, although it is not obvious why it should be more pronounced at low than at high ionic strength. (3) Since the reactants are polyelectrolytes, their average conformations in solution are dependent on the ionic strength, and it is conceivable that the rate of formation of the helical complex is in part determined by the conformations of the reactants.

Rich[13] concluded on the basis of measurements of the rate of combination of poly A and polyinosinic acid to form a two-stranded helical complex that this process follows first order kinetics. However, since he employed conventional spectrophotometry, his first observation could not be made until the reaction was 40 per cent complete. Some of our experiments at low ionic strength would appear to follow first order kinetics with reasonable accuracy if the observations during the first half of the reaction were omitted.

The approach of the rate to zero at the melting temperature of the helix is quite analogous to the phenomenon of the "ceiling temperature" [14] in addition poly-

merizations. The usually accepted explanation of this phenomenon is that the rates of both the propagation and depolymerization processes increase with temperature, with the activation energy for depolymerization being larger than that for propagation, so that at a sufficiently elevated temperature the two rates become equal and no reaction takes place. Similarly, in the present case, at approximately the helix-melting temperature the rate of growth of nuclei and the rate of disordering helical units become equal.

Summary.—The kinetics of formation of the double-stranded helical complex poly (A + U) has been studied by the spectrophotometric stopped-flow technique. The reaction follows complex kinetics, initially of second order form, then deviating toward first order. The initial rate is first order with respect to each polymer component, and is strongly dependent on the ionic strength in the direction expected for a reaction between ions carrying charges of the same sign. The reaction rate decreases with increasing temperature, approaching zero at the helix melting temperature.

The authors are indebted to Dr. Alexander Rich for the polymer samples and their characterization, and to Dr. Bruno Zimm for several stimulating discussions. This research was supported in part by grants from the United States Public Health Service (RG-4725) and the National Science Foundation (G-9625).

* Contribution No. 1638 from the Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut.

† Postdoctoral Fellow of the Division of General Medical Sciences, U. S. Public Health Service.

[1] Warner, R. C., *Federation Proc.*, **15**, 379 (1956).

[2] Rich, A., and D. R. Davies, *J. Am. Chem. Soc.*, **78**, 3548 (1956).

[3] Felsenfeld, G., and A. Rich, *Biochim. Biophys. Acta*, **26**, 457 (1957).

[4] Fresco, J. R., and P. Doty, *J. Am. Chem. Soc.*, **79**, 3928 (1957).

[5] Chance, B., in *Investigations of Rates and Mechanisms of Reactions*, ed. S. L. Friess and A. Weissberger, (New York: Interscience Publishing Co., 1953), p. 690.

[6] Spencer, T., and J. M. Sturtevant, *J. Am. Chem. Soc.*, **81**, 1874 (1959).

[7] Beers, R. F., Jr., and R. F. Steiner, *Nature*, **179**, 1076 (1957).

[8] Fresco, J. R., and B. M. Alberts, these PROCEEDINGS, **46**, 311 (1960).

[9] See, for example, Doty *et al.*, these PROCEEDING, **45**, 482 (1959).

[10] The observation by Steiner and Beers[11] of changes in light scattering continuing long after completion of the hypochromic shift shows that the use of absorbancy as a measure of the extent of reaction must be restricted to the first few minutes after initiation of the reaction under the condition employed in our experiments.

[11] Steiner, R. F., and R. F. Beers, Jr., *Biochim. Biophys. Acta*, **33**, 470 (1959).

[12] Felsenfeld, G., *Biochim. Biophys. Acta*, **29**, 133 (1958).

[13] Rich, A., *Inst. Internat. Chim. Solvay*, Brussels, June, 1959.

[14] Dainton, F. S., and K. J. Ivin, *Nature*, **162**, 705 (1948).

THE CRYSTAL STRUCTURE OF TERRAMYCIN HYDROCHLORIDE

By Y. Takeuchi and M. J. Buerger

UNIVERSITY OF TOKYO AND MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Communicated July 25, 1960

The problem of determining the arrangements of atoms in crystals directly from X-ray diffraction data has engaged the attention of crystallographers for some time, and a number of approaches to this general problem have been presented. We have been especially interested in the approach which makes use of image-seeking functions in solving the Patterson synthesis,[1] which, in turn, is based entirely on observable X-ray diffraction data. Many simple and moderately complex structures have been solved directly from X-ray diffraction data by this route. We wished to test the power of the method on a somewhat more complex structure, and we regarded the antibiotics as well suited to this purpose.

In order to focus attention on the problem and not on the details of gathering the X-ray diffraction data, we canvassed our colleagues for a full set of three-dimensional diffraction data for a structure as yet unsolved. Professor Ray Pepinsky kindly made available to us such data for terramycin hydrochloride which had been gathered in his laboratory by Dr. J. H. Robertson. Professor Pepinsky further supplied us with sections of the three-dimensional Patterson function for these data, as computed by his X-RAC.

Space group and unit cell: Crystals of terramycin hydrochloride are orthorhombic, space group $P2_12_12_1$, and have the following cell dimensions:

$$a = 11.19\text{Å}$$
$$b = 12.49$$
$$c = 15.68$$

This cell contains 4 molecules of terramycin hydrochloride, which has been assigned the formula[2] $C_{22}H_{24}N_2O_9 \cdot HCl$. Since space group $P2_12_12_1$ has only general fourfold positions, it is evident that the terramycin hydrochloride molecule can be regarded as the asymmetric unit.

Intensity data: The diffraction intensities had been gathered by Dr. J. H. Robertson from Weissenberg photographs made with copper radiation. The intensities of some 2,500 reflections had been measured and corrected for Lorentz and polarization factors. These were substantially all the reflections within the recording range of copper radiation.

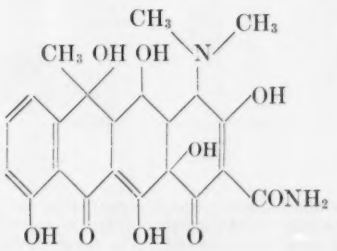
Patterson synthesis: The Patterson function was computed by X-RAC in the form of sections parallel to (001) and spaced at intervals of $c/60$. X-RAC presents these sections in the form of the contoured function, which is photographed as it appears on a kinescope screen. Since sections made in this way contain distortion, the 30 sections of the asymmetric unit were redrawn for the more exacting requirements of image-seeking functions.

Image-seeking functions: Image-seeking functions provide an approximate picture of the electron density of the structure. Their use for space group $P2_12_12_1$ has been specifically discussed.[3] The action of the function is briefly as follows.

Each atom (the somewhat heavy chlorines in particular, in this case) contributes to the Patterson function an image of the electron density of the structure, as seen from itself. If there are n such atoms, the Patterson function contains n overlapping images of the structure, and the image-seeking function brings these together. This is well done by the particular image-seeking function known as the minimum function. Such individual functions can be combined to give more powerful functions, and, with 4 symmetry-related chlorine atoms per cell, the fairly powerful M_4 function can be constructed.

To start such a set of functions, the locations of the somewhat heavy chlorine atoms are required. These can be found in the usual way[1] from the Patterson function itself. In terramycin hydrochloride, a representative atom is found to have fractional coordinates 0.075, 0.080, 0.000. The occurrence of the atoms at level $z = 0$ is an unfortunate specialization, since it introduces a false mirror parallel to (001) into the minimum function based upon the atom. Because of this, instead of providing a normal solution to the structure, the M_4 function in this case provides the structure plus its enantiomorphic image as reflected in a mirror parallel to (001). This prevents direct study of the shape of the molecule since both the molecule and its undesired reflection are seen interlaced in space. The only unambiguous information which can be obtained is the shape of the molecule as seen projected on the (001) plane, since in this projection the molecule and its enantiomorphic reflection coincide.

Structure determination: Because of the unfortunate degeneracy of the M_4 function, just mentioned, we were prevented from utilizing the minimum function in a straightforward fashion in three dimensions, and we were forced to use comparatively weak methods. Specifically, we attempted to recognize the shape of the molecule from the M_4 function, both as projected on (001). The preferred shape of the molecule from chemical evidence is given by Hochstein *et al.*:[4]



We failed to find a good fit for such a molecule in our projection of M_4, and so we discontinued our work temporarily.

Meanwhile, Hirokawa *et al.*[5] have determined the structure of aureomycin hydrochloride, and this revealed why we failed to fit the chemically expected molecule to our M_4 function. Their evidence makes it plain that we had greatly underestimated the effects of thermal agitation. The atoms of aureomycin hydrochloride associated with ring IV have a very high thermal agitation. It was some of the atoms of the corresponding ring of terramycin which appeared to be missing in our M_4 function.

If this large thermal motion of the atoms of ring IV were taken into account, it was plain that we could fit the shape of the preferred chemical terramycin molecule to our M_4 projection. Figure 1 shows the fit so obtained. Peaks A, B, C, D, E, and F are heavy and, accordingly, must be either the heavier oxygen or nitrogen atoms. Peak A is about 3.1 Å from a chlorine atom, and is doubtless the nitrogen atom of the strongly basic dimethylamino group. Peaks for $CH_3^{I}(4)$ and $CH_3^{II}(4)$ do not appear in the M_4 projection. In our earlier trial we had tried this location for the molecule many times but always discarded it because we could not find these atoms.

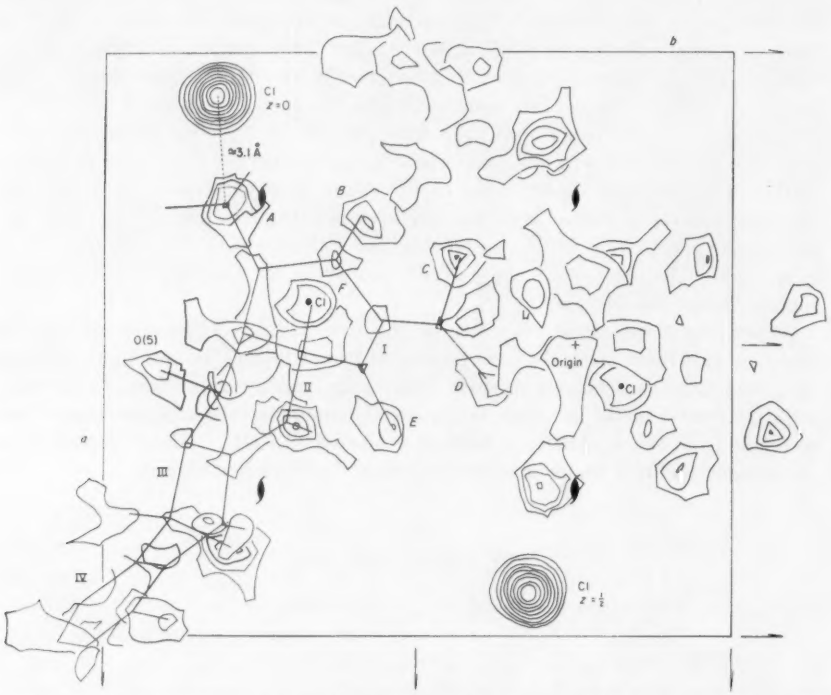


FIG. 1.—The peaks of the minimum function $M_4(xyz)$ for terramycin hydrochloride projected on (001), but with peaks due to false enantiomorphic molecule omitted, and two overlapping chlorine atoms (indicated by heavy dots) also omitted. The range of the function corresponds to the bounded projection $\int_0^{1/2} \rho(xyz)\, cdz$. The customary origin of space-group $P2_12_12_1$ is indicated. The terramycin molecule is outlined, and its four rings are numbered.

We now computed the amplitudes of the $0kl$ and $h0l$ reflections for this placement of the molecule, and were delighted to find very promising agreement with our measured intensities. The disagreement factor, R, for all these reflections was only about 31 per cent, the coordinates being based only on the M_4 peaks.

Encouraged by this, we made a projection of the electron density $\rho(yz)$ for the model. This is shown in Figure 2. The result is very satisfactory, for the series converged well, and there are no false peaks. We feel, therefore, that we have

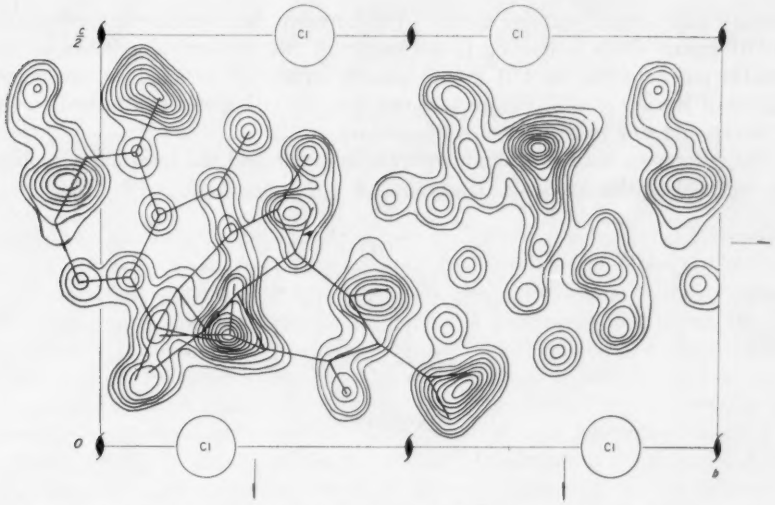


FIG. 2.—Electron-density projection $\rho(yz)$, of terramycin hydrochloride, based upon phases determined from the minimum function, $M_4(xyz)$, shown in Fig. 1. The origin is taken on a twofold screw axis for convenience of computation.

determined the structure of terramycin hydrochloride and know the structure with limited precision. The coordinates of the atoms of this structure are given in Table 1.

TABLE 1

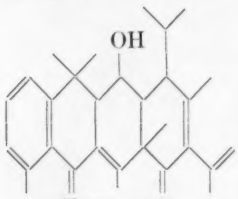
COORDINATES OF ATOMS IN TERRAMYCIN HYDROCHLORIDE

	x	y	z		x	y	z
Cl^-	0.075	0.080	0.000	$C_{amide}(2)$	−0.030	0.786	0.078
O(1)	.130	.705	.182	C(3)	− .146	.618	.099
$O_{amide}(2)$	− .148	.810	.036	C(4)	− .13	.51	.12
O(3)	− .225	.643	.057	C(4a)	− .024	.472	.188
O(5)	− .041	.338	.093	C(5)	.082	.437	.130
O(6)	.340	.459	.122	C(5a)	.175	.378	.192
O(10)	.474	.359	.433	C(6)	.295	.360	.145
O(11)	.320	.491	.384	C(6a)	.37	.30	.20
O(12)	.139	.590	.357	C(7)	.44	.22	.19
O(12a)	− .071	.584	.291	C(8)	.51	.18	.26
N(4)	− .243	.459	.143	C(9)	.52	.28	.33
$N_{amide}(2)$	.045	.843	.081	C(10)	.45	.31	.36
$CH_3(4)$-I	− .29	.47	.22	C(10a)	.38	.35	.29
$CH_3(4)$-II	− .23	.35	.14	C(11)	.325	.450	.315
$CH_3(6)$	.28	.30	.06	C(11a)	.18	.46	.267
C(1)	.035	.656	.180	C(12)	.125	.545	.284
C(2)	−0.48	0.694	0.122	(12a)	0.02	0.56	0.23

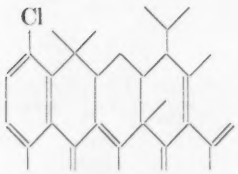
Discussion of the Structure.—We are proceeding with the refinement of the structure, but meanwhile some general remarks can be made about it. There is a close similarity between the terramycin molecule and the aureomycin molecule as already suspected on chemical grounds. But there are some notable differences. While the aureomycin molecule has a chlorine atom, this is not a part of the terramycin molecule. The bond distance from the Cl of terramycin hydrochloride to peak *A* in Figure 1 is about 3.1 Å. Peak *A* is interpreted as a nitrogen atom of the

strongly basic dimethylamino group. Furthermore, the terramycin molecule has an OH group which is missing in aureomycin. Stereochemically, there are two possible positions for the OH group, namely equatorial or axial, in the nomenclature of Barton *et al.*[6] Figure 1 shows that the OH group is attached to C(5) of the second ring in equatorial configuration.

The differences between the aureomycin molecule and the terramycin molecule are outlined in the following diagrams.



Terramycin



Aureomycin

[1] Buerger, Martin J., *"Vector Space and Its Application in Crystal-Structure Investigation,"* (New York: John Wiley and Sons, 1959).

[2] Regna, Peter P., and I. A. Solomans, "The chemical and physical properties of terramycin," *Ann. N. Y. Acad. Sci.*, **53,** 229–237 (1950).

[3] Hahn, Theodor, and M. J. Buerger, "The crystal structure of diglycine hydrochloride, $2(C_2H_5O_2N),HCl$," *Z. Krist.*, **108,** 419–457 (1957).

[4] Hochstein, F. A., C. R. Stephens, L. H. Conover, P. P. Regna, R. Pasternack, K. J. Brunings, and R. B. Woodward, "Terramycin, VII: The structure of terramycin," *J. Am. Chem. Soc.*, **74,** 3708 (1952).

[5] Hirokawa, S., Y. Okaya, F. M. Lovell, and R. Pepinsky, "The crystal structure of aureomycin hydrochloride," *Z. Krist.*, **112,** 439–464 (1959).

[6] Barton, D. H. R., O. Hassel, K. S. Pitzer, and V. Prelog, "Nomenclature of cyclohexane bonds," *Nature*, **172,** 1096–1097 (1953).

POPULATION STUDIES IN PREDOMINANTLY SELF-POLLINATED SPECIES, I. EVIDENCE FOR HETEROZYGOTE ADVANTAGE IN A CLOSED POPULATION OF BARLEY

By S. K. Jain and R. W. Allard

UNIVERSITY OF CALIFORNIA, DAVIS

Communicated by G. Ledyard Stebbins, August 18, 1960

Breeding systems play an important role in determining the genetic structure and evolutionary dynamics of populations. In flowering plants the mechanisms governing mating systems show great diversity. For the evolutionist the important distinction among these systems is the degree of inbreeding versus outbreeding they produce. Three modal classes can be recognized: exclusive or predominant outcrossing, predominant selfing, and mixed selfing and outcrossing. Mather[1] emphasized that each class is likely to have a characteristic way of meeting the necessary compromise between the requirements of high immediate fitness and long-term flexibility. Many of the evolutionary properties of the so-called Mendelian populations of outbreeders, such as balanced polymorphisms, coadaptation, and genetic homeostasis, are associated with the storage of variability in heterozygotes.[2] In self-fertilizing groups, in contrast, the store of variability is postulated to be in the form of a genetically diverse array of sympatric homozygous lines that occasionally intercross to produce, on subsequent selfing, a series of new homozygous lines.[3] Thus the population is presumed to consist primarily of homozygous subunits isolated, not spatially but by their breeding habit, and secondarily of a small proportion of recently formed heterozygotes. The present study was undertaken to investigate the level of heterozygosity in an experimental population of barley, a predominantly self-fertilizing organism, and the adaptive significance of heterozygotes in the storage and release of genetic variability.

The experimental material was Composite Cross V, a population developed by the late H. V. Harlan and associates from intercrosses among 31 varieties of barley. In 1937 the varieties were first crossed in 15 pairs. During the next three years the F_1 hybrids (plus the uncrossed parent) of the preceding cycle were again crossed to produce ultimately a single hybrid stock. The first selfed generation, designated F_2, was grown in 1941, the F_3 in 1942, and so on, from random samples of seeds grown without conscious selection according to usual agricultural practice. We are indebted to C. A. Suneson[4] (who grew the F_3-F_{18} generations at Davis, California from 1942 subsequently) for making available seed of the F_3, F_5, F_{13}, F_{15}, and F_{18} generations of this composite cross. The following loci were chosen for study: *B/b*, black versus white pericarp color; *S/s*, long-haired versus short-haired rachilla; *G/g*, dentate versus nondentate lemma veins; Bl_1/bl_1—Bl_2/bl_2, blue versus white aleurone color (complementary genes on different chromosomes); *Sh/sh*, erect versus prostrate early-growth habit; *R/r*, rough versus smooth awns, *Bt/bt*, brittle versus tough rachis; *E/e*, long versus short awn on the glumes.

The plan of the experiment was, first, to determine changes in frequency of homozygotes and heterozygotes at the several marker loci over generations, and then to identify the forces responsible for the observed changes by comparison with

a genetic model taking into account the major factors that are recognized as affecting genotypic frequencies. These factors fall into two broad categories; (*a*) directed processes, including mutation, migration, selection, and the mating system, and (*b*) random processes, including fluctuations in the directed processes, and sampling accidents associated with population size. Since this population was closed and large (10,000–15,000 individuals/generation), migration and random drift presumably had only minor effects on changes in gene and genotypic frequencies. Further, it is doubtful that mutation had much effect over the 18-generation period involved. Hence, among the major directed and random evolutionary forces, the ones most likely to have produced significant directional changes in genotypic frequencies in the population appear to have been natural selection acting on unequal adaptive values of the three genotypes at a given locus, and inbreeding. The major emphasis was therefore on determining the breeding system of the population, particularly the amount of selfing versus outcrossing, and, by means of a census, determining changes in genotypic frequencies (these estimates permit inferences concerning the selective values of the three possible genotypes at each marker locus).

TABLE 1

ESTIMATED AMOUNTS OF OUTCROSSING (t) IN PER CENT

Locus	Generation and year grown	Number of progenies	Number of outcrosses	Outcrossing (t)	σ_t
B/b	F_3 (1942)	1541	2	1.04	0.0073
Bl/Bl	F_4 (1959)	3162	18	0.94	0.0023
	F_{19} (1959)	2335	18	1.28	0.0031
S/s	F_3 (1942)	535	3	2.35	0.0134
	F_5 (1944)	678	3	2.20	0.0125
	F_{15} (1954)	372	0	0.00	0.0000
	F_{18} (1957)	378	1	1.00	0.0099
R/r	F_3 (1942)	225	0	0.00	0.0000
	F_{18} (1957)	446	1	0.42	0.0042

It should be emphasized that the hereditary units in consideration here are undoubtedly not single loci but chromosome segments marked by the loci studied. In this barley population, synthesized by intermating 31 homozygous parents followed by predominant self-fertilization, the mean length of segment per chromosome can be computed[5] to be 40 or more crossover units on the average at the end of the intercrossing phase, and 20 to 30 crossover units after 18 generations of selfing. For convenience, however, these segments are herein designated by the genetic symbol of the marker gene.

Results.—The mating system in this population was assessed by counting the number of dominant phenotypes among the progenies of a random sample of recessive genotypes for four of the marker loci. Following a method suggested by Fyfe and Bailey[6] the data yielded the maximum likelihood estimates of the amount of outcrossing given in Table 1. Since there was no consistent difference between readings in different years or generations, the amount of outcrossing in this population appears to have varied within a narrow range about a mean of approximately 1 to 2 per cent.

The proportions of three genotypes, two homozygotes, and the heterozygote in a given generation were estimated from progeny tests of random samples of plants

drawn from the population. These estimates are given in Table 2, in which P, Q, and R respectively denote the dominant homozygote, recessive homozygote, and heterozygote. The same data are shown in graphical form in Figure 1.

From the values of P, Q, and R it is clear that genotypic frequencies changed markedly for certain of the 8 loci during 18 generations of natural selection. The frequency of the dominant phenotype (i.e., P + R) increased sharply for *R*, Bl_1, and Bl_2, decreased for *B* and *Bt*, and remained much the same for *E* and also *S* in later generations. It should be noted that changes in the proportion of P and Q can lead to fixation of one or the other allele. Such changes are readily apparent in Figure 1, in which *B*, *E*, *R*, *G* and *Bl* have moved toward either of the laterals marked as P or Q. Only locus *B*, however, approaches fixation by the F_{18} genera-

TABLE 2

ESTIMATES OF GENOTYPIC FREQUENCIES (P, Q, R) AND SELECTION COEFFICIENTS (x, y)

Locus	Genotype	F_2	F_5	s = 1 x	s = 1 y	s = 0.98 x	s = 0.98 y	F_{12}	F_{15}	F_{18}	s = 1 x	s = 1 y
B	P	0.059	0.074	1.00	1.25	1.06	1.31	0.013	0.002	0.000	1.00	1.25
	Q	0.810	0.906	...	...	...	...	0.987	0.998	1.000	...	...
	R	0.131	0.020	...	...	...	...	0.000	0.000	0.000	...	...
S	P	0.072	0.148	...	...	...	...	0.221	0.226	0.261	...	...
	Q	0.594	0.745	0.75	0.90	0.81	0.96	0.757	0.739	0.731	1.02	0.99
	R	0.334	0.107	...	...	...	...	0.022	0.035	0.008	...	...
G	P	0.357	0.556	...	...	...	...	0.575	0.657	0.671	...	...
	Q	0.357	0.350	0.95	0.75	1.04	0.82	0.396	0.291	0.308	0.80	0.80
	R	0.286	0.094	...	...	...	...	0.029	0.052	0.021	...	...
E	P	0.349	0.270	...	...	...	...	0.304	...	0.248	...	...
	Q	0.535	0.616	0.41	0.52	0.47	0.59	0.636	...	0.693	...	...
	R	0.116	0.114	...	...	...	...	0.060	...	0.059	...	...
*Bl_1 or Bl_2	P	0.408	0.487	...	...	...	...	0.505	0.548	0.613	...	...
	Q	0.293	0.304	0.56	0.50	0.61	0.54	0.296	0.280	0.255	0.52	0.47
	R	0.299	0.209	...	...	...	...	0.199	0.172	0.132	...	...
R	P	0.266	0.370	...	...	...	...	...	...	0.489	...	...
	Q	0.614	0.557	0.70	0.59	0.82	0.68	...	...	0.411	...	...
	R	0.120	0.073	...	...	...	...	...	...	0.100	...	...
Bt	P	0.090	0.122	...	...	...	...	...	...	0.123	...	...
	Q	0.758	0.835	0.90	1.00	0.96	1.06	...	...	0.870	...	...
	R	0.152	0.043	...	...	...	...	...	...	0.007	...	...
Sh	P	0.544	0.615	...	...	...	...	...	...	0.682	...	...
	Q	0.275	0.276	0.65	0.58	0.71	0.63	...	...	0.290	...	...
	R	0.181	0.109	...	...	...	...	...	...	0.028	...	...

x and y of columns 5–8 were estimated by minimizing chi-square of deviation between the observed and expected frequencies in F_5, starting with those observed in F_2. x and y of columns 12 and 13 were obtained by fitting to the data of F_{15}-F_{18} interval.

* The frequency of single heterozygotes (i.e., those segregating as 3:1) was assumed to be the same at both loci, so that P, Q, and R would also become identical.

tion. Loci *S* and *Bt*, in contrast, seem close to an equilibrium state, in which the two alternative alleles at each locus may continue to coexist in the population for many additional generations. A necessary and sufficient condition for stability of an equilibrium involving only the homozygotes is simply equality of their adaptive values.

It is also clear from the values of *R* that the rate of decay of heterozygosity differs for these 8 loci. At loci *B*, *Bt*, *G*, and *S*, the proportion of heterozygotes (R) decreased steadily at about the rate expected, with 1 to 2% of outcrossing. The decrease was slower at loci Bl_1 (or Bl_2) and *Sh*, but in both instances seemed to approach the equilibrium parabola steadily. The level of heterozygosity at loci

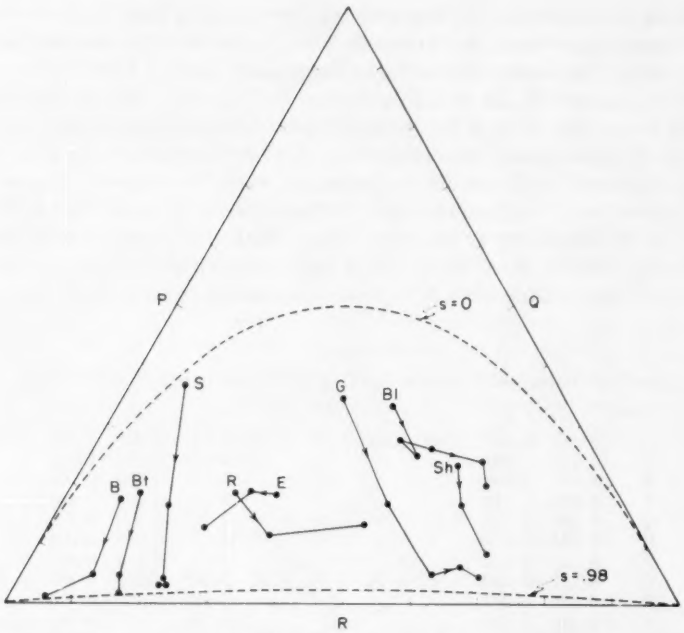


FIG. 1.—Trilinear diagram showing the genotypic frequencies (P, Q, R) given in Table 1. The perpendicular distances from the three sides marked P, Q, and R give the proportions of the three genotypes at the loci indicated; each point corresponds to a generation and the arrows point out the direction of change. The broken lines represent equilibrium parabolas, in absence of selection, for $s = 0$, and $s = 0.98$.

R and *E*, in contrast, decreased only slightly in the early generations and then seemed to become stabilized at an equilibrium value higher than expected for no selection and 98 to 99 per cent of selfing. If, as deduced earlier, natural selection is the primary force acting in differential survival of genotypes, it operates with different intensities on the various chromosome segments carrying marker genes.

Quantitative study of the effects of selection is difficult because the data do not provide a direct measure of the adaptive values for the three genotypes at any given locus. Even so, inferences about adaptive values can be made by comparing observed genotypic frequencies with the frequencies expected for given levels of selfing and outcrossing. Hayman,[7] taking advantage of the fact that populations such as that under consideration consist of two components, one that selfs and one that outcrosses (presumably at random), derived a set of proportionalities that express changes in genotypic frequencies at a locus. In Hayman's model, P_n, Q_n, and R_n ($P_n + Q_n + R_n = 1$) respectively denote the frequencies of genotypes *AA*, *aa*, and *Aa* at a locus *A*/*a* in the n^{th} generation; s and t ($s + t = 1$) represent the relative amounts of selfing and random outcrossing; and the adaptive values of *AA*, *aa*, and *Aa* are assumed to be in the ratio of x:y:1. If the population is large enough that random fluctuations in P_n, Q_n, and R_n are trivial, the genotypic frequencies become

$$\left.\begin{aligned} P_n + 1 &\propto x\{s(P_n + {}^1/_4R_n) + t(P_n + {}^1/_2R_n)^2\} = x\{P_n + {}^1/_4(R_n - te_n)\} \\ Q_n + 1 &\propto y\{s(Q_n + {}^1/_4R_n) + t(Q_n + {}^1/_2R_n)^2\} = y\{P_n + {}^1/_4(R_n - te_n)\} \\ R_n + 1 &\propto {}^1/_2sR_n + 2t(P_n + {}^1/_2R_n)(Q_n + {}^1/_2R_n) = R_n - {}^1/_2(sR_n - te_n) \end{aligned}\right\} (1)$$

in which $e_n = 4P_nQ_n - R^2_n$. These proportionalities become equations when $x = y = 1$. When P, Q, R in various generations, and s and t, are known, the adaptive values x and y can be determined uniquely from the proportionalities (1) by successive iterations until the set is found that gives minimum χ^2.

Since the magnitude of the selective forces might not have been the same from F_3 to F_{18}, x and y were estimated for the intervals F_3–F_5 and F_{15}–F_{18}. The estimates were made for $s = 1$ for both intervals and for $s = 0.98$ for the early generation interval to determine the effect of 2% outcrossing on the adaptive values. These estimates are given in Table 2, together with observed values of P, Q, and R. A point to note is that calculations based on the observed rate of outcrossing change the estimates of x and y only slightly toward higher values. The observed values of x and y indicate that adaptive values were usually not equal for the three genotypes at any given locus. Thus, for loci *E* and *Bl*$_1$ (or *Bl*$_2$), the homozygotes apparently contribute to the next generation only about half as many progeny as does the heterozygote. The contribution of the heterozygote may also have been substantially greater at loci *Sh*, *R*, and *S*. Homozygotes and heterozygotes appeared to be about equal in fecundity at the remaining loci except *B*, at which locus there may have been selection against the heterozygote.

Interpretation of Results.—The most important fact that emerged from this study is that the proportion of heterozygotes at a number of marker loci did not decrease at the rate expected under the observed levels of selfing and outcrossing. Thus, although outcrossing is capable of maintaining a few heterozygotes in the population (see also the equilibrium parabola for $s = 0.98$ in Fig. 1), heterozygote advantage needs to be invoked to account for the level of heterozygosis observed at most loci.

The most obvious explanation for the advantage of heterozygotes is overdominance (in terms of fitness) associated with the marker loci themselves. In several reports in the literature, claims have been made for single-locus overdominance. Although true overdominance cannot be ruled out here, neither can it be demonstrated; the hereditary units in consideration are chromosome segments likely to be of the order of 40 or more crossover units long at the end of the intercrossing phase of the experiment, and 20 to 30 after 18 generations of predominant selfing.

Striking overdominance associated with chromosome segments could result from any or a combination of several genetic causes. For example, a theoretical study of selection on two linked loci has shown that linkages can cause the effects of selection to be spread over a chromosome segment and that the levels of heterozygosity observed can be accounted for on the basis of less overdominance at two or more loci under selection than need be assumed for any single locus.[8] Pseudo-overdominance caused by linkage disequilibrium with excess of repulsion phase linkages, or, alternatively, the summed effect of several linked loci each partially dominant, could produce the same result. Linkage effects in this population seem even more important because of the fact that parallel studies of quantitative characters in this population indicate that selection favors intermediate phenotypes,

which might be associated with a preponderance of repulsion phase linkages.[8] Whatever its genetic causes, the existence of striking heterozygote advantage in a predominantly self-fertilizing population raises the question of its significance in the immediate fitness and the long-term future of the population.

Large evolutionary potential is presumably a feature of composite crosses or other populations derived from intercrosses among numerous diverse strains. Intermating followed by segregation under predominant self-fertilization makes many new combinations of genes available, offering great scope for adaptive improvements. Wright[9] suggested that this can result in a large number of adaptive peaks; the persistence into the eighteenth generation of a large proportion of the 2^8 possible phenotypic combinations possible for the 8 marker loci indicates that this happened in the present population.[8] The situation is in sharp contrast to that of experimental populations made up of a mixture of homozygous lines; there the number of adaptive peaks is small, as shown by Harlan and Martini,[10] and the genetic potential of the population is limited as a result of the rapidity with which one or a few genotypes becomes predominant.

The maintenance of genetic variability in Composite Cross V appears to depend not only on the low incidence of intercrossing among the many diverse genotypes contained in the population, but also on the substantial heterozygote advantage, which prolongs the period over which segregation can take place. Thus heterozygote advantage is equivalent in effect to an increase in the rate of outcrossing.

These results in barley have an important bearing on our understanding of genetic variability in "self-fertilizing" species. Rarely if ever is selfing complete, a fact that has apparently not always been recognized. Although a self-fertilizing population must clearly be at a disadvantage with respect to flexibility unless hybridization occurs between different genotypes in the population, equally, excessive crossing may lead to undue hybridity. Thus, according to Darlington and Mather,[11] each population has an optimum degree of hybridity to which it is accustomed and departures from this optimum, no matter in what direction, are followed by unbalance in the genetic system. Control of segregation and recombination through such devices as regulation of chiasma frequencies, together with control of the rate of outcrossing are postulated to be involved in the achievement of optimum hybridity. The present results indicate that level of heterozygote advantage is another interrelated factor. Indeed, it could be argued that high heterozygote advantage has evolved in self-pollinated species as a substitute for high levels of outcrossing. This is clearly the opposite of Lewontin's[12] view that "no claim for greater fitness of heterozygotes is made in the case of self-fertilizing species, for such organisms have clearly foregone the advantages of heterozygosity in the course of realizing their particular evolutionary mode."

The observation of striking heterozygote advantage in self-fertilizing species thus suggests two modifications of the classical concepts of population structure in such species. First, the adoption of a mating system involving predominant self-fertilization is not necessarily accompanied by relinquishment of the advantages of heterozygosity. Indeed, a few highly fit heterozygotes are capable of endowing a population with the same mean fitness as a larger proportion of mildly superior heterozygotes. Second, the store of variability in self-fertilizing species apparently is not entirely associated with genetic diversity among multiple homozygous co-

existing familial lines, since the storage potential of even a few superior heterozygotes could be substantial.

It is possible that the observed heterozygote advantage is a transient property that, after contributing to genetic diversity following hybridization, disappears at equilibrium. The population studied here, in common with most artificial populations, is clearly not at equilibrium at most loci; hence, issues such as this are best studied in natural populations.

Finally, it should be emphasized that the foregoing analysis of changes at marker loci over generations was based on deterministic models although the actual processes must certainly have been stochastic. During the 18 years Composite Cross V was propagated, it doubtless encountered a succession of different environments that could well have caused fluctuations in the selective values of different genotypes. Other factors, such as variations in population density and changes in composition of the population over time, could also have caused selective values to vary. For example, comparisons of x and y for early and late generations suggest that heterozygote advantage disappeared at the S locus as heterozygotes became infrequent, whereas at the G locus the opposite change seemed to occur. Although further studies will be required to determine the consistency and possible significance of such changes in adaptive values, Monte Carlo investigations on a digital computer of a number of cases simulating the present experiment have suggested that changes in heterozygote frequencies in populations as large as the present one are not much influenced by the random elements (amount of outcrossing, adaptive values; unpublished data of Allard, Baker, and Baker).

Summary.—Analysis of a barley population in terms of changes in genotypic frequencies occurring over 18 generations, under a mating system involving predominant self-fertilization, indicated that decay in heterozygosity was slower at several marker loci than expected for the observed amount of outcrossing. These results were explained on the basis of heterozygote advantage associated with segments of chromosomes. It was postulated that heterozygote advantage may have important adaptive implications in the maintenance of population variability, and hence in the evolutionary potential of self-fertilizing species.

[1] Mather, K., *Biol. Rev.*, **18**,32–64(1943).
[2] Dobzhansky, Th., *Cold Spring Harbor Symp. Quant. Biol.*, **20**, 1–15(1955).
[3] Stebbins, G. L., *Am. Naturalist*, **91**, 337–354(1957).
[4] Suneson, C. A., *Agron. Jour.*, **48**, 188–191(1956).
[5] Hanson, W. D., *Genetics*, **44**, 857–868(1959).
[6] Fyfe, J. L., and N. T. J. Bailey, *Jour. Agric. Sci.*, **41**, 371–378(1951).
[7] Hayman, B. I., *Heredity*, **7**, 185–192(1953).
[8] Jain, S. K., Ph.D. dissertation, University of California, (1960).
[9] Wright, S., *Am. Naturalist*, **90**, 5–24(1956).
[10] Harlan, H. V., and M. L. Martini, *Jour. Agric. Res.*, **57**, 189–199(1938).
[11] Darlington, C. D., and K. Mather, *The Elements of Genetics*, (London: Allen Unwin, 1952). p. 446.
[12] Lewontin, R. C., *Evolution*, **12**, 494–503(1958).

*GENETIC CONTROL OF SERUM TRANSFERRIN TYPE IN MICE**

D. C. SHREFFLER†

DIVISION OF BIOLOGY, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated by George W. Beadle, August 22, 1960

Relationships between genes and proteins are being defined in detail through genetic studies of the fine structure of the chromosome and chemical analyses of the molecular structures of genetically variant proteins. The most extensive chromosomal fine structure investigation has been done with bacteriophage.[1] The best evidence concerning the effects of mutation on protein structure has come from chemical and genetic studies of human hemoglobins.[2] Eventually genetic fine structure studies and protein structure analyses will be combined for given genetic loci and the proteins they control. The most elegant studies of this type will probably be done with microorganisms, because of their advantages for genetic study. It is important, however, also to extend our understanding of gene-protein relationships in higher animals, particularly mammals. The hemoglobin work has demonstrated that such investigations are possible in mammals. It should be worthwhile to analyze other protein variant systems in the same way, in order better to understand the detailed mechanisms and relationships involved in the genetic control of a variety of different proteins. Such studies require variant systems which can be readily isolated and purified, which have been shown to be genetically determined, and which, preferably, occur in species adapted to intensive genetic study.

A genetic difference has been observed in the β-lactoglobulins of cow's milk,[3] and a number of physicochemical studies of these differing proteins have been carried out.[4–6] Genetically determined antigenic differences in the γ-globulins of rabbits have been studied by gel diffusion and immunoelectrophoresis.[7–9] Chemical investigations of rabbit γ-globulins have been in progress for many years. Investigations of fractions obtained by papain digestion by Porter[10] and by peptic digestion by Nisonoff *et al.*[11] offer hope for identification of the antigenic differences with specific portions of the molecule. Observations with starch gel electrophoresis of a number of genetically determined serum protein variant systems in humans (reviewed by Smithies[12]), cattle,[13, 14] sheep,[14] goats,[14] horses,[15] and swine[16] have provided materials which may prove useful for structural studies. Adequate techniques are available so that many of these serum proteins could be isolated and purified without serious difficulty. However, with all of the species in which serum protein variants have been identified, as well as with the human hemoglobin studies, there is difficulty in conducting intensive genetic studies of the sort which would be desirable for the most precise definitions of gene-protein relationships in mammals, in terms of allelism, linkage, intragenic recombination, etc.

The most favorable mammalian species available for such investigations is the mouse. Well-defined linkage groups, adequate markers, the availability of highly inbred lines, and a short generation interval make it very useful for genetic study. Hemoglobin variants have been observed in the mouse and have been well studied genetically.[17] However, no other well-defined protein variants have been reported in the mouse. Thompson *et al.*[18] have described an apparent genetic difference

between certain lines of mice in quantity of β_1-globulin, determined by moving boundary electrophoresis, but this might involve only an increase in some lines of a component common to all, rather than a molecular difference. The genetic evidence is based only on F_1's; segregating progeny could not be tested because it was necessary to pool sera from a large number of mice for each moving boundary run.

To detect useful protein variant systems in mice, sera from a number of inbred lines have been screened, utilizing two of the most powerful techniques presently available for detection of molecular differences in proteins—starch gel electrophoresis and the Ouchterlony gel diffusion technique. Presented here are the results of a genetic study of one of several serum protein differences which have been detected, an electrophoretically distinguishable serum transferrin variant.

Materials and Methods.—Thirteen inbred lines of mice were surveyed: A/Jax, A/He, AKR, BALB/c, CBA, C3H, C57Bl/6, C57Br, DBA/2, Flexed, RFM, WB, and WC. For genetic study of the transferrin difference, reciprocal crosses were made between CBA and C57Bl/6 (B/6) to produce F_1 progeny. $F_1 \times F_1$ crosses and reciprocal backcrosses to each parental line were made with both groups of F_1's. Progeny were weaned and individually identified at four to five weeks, but no mouse was tested for serum protein type until it had reached at least eight weeks of age. Mouse sera exhibit a number of qualitative and quantitative changes from birth to three to four weeks of age;[19] therefore testing was delayed until it was certain that only adult characteristics would be observed.

Blood samples were collected from the ventral tail vein in silicone-treated tubes. Sera were tested fresh or after storage at $-10°C$. Only slight changes in electrophoretic patterns of the sera were noted after freezing.

The serum components were separated by horizontal starch gel electrophoresis.[20] Gels were prepared with approximately 13 gm. hydrolyzed starch (Connaught Medical Research Laboratories, Toronto, Canada) per 100 ml borate buffer, pH 8.5, approximately 0.02 molar in borate. Exact starch and buffer concentrations varied with different lots of starch. Fifty μl of each serum sample were pipetted, without supporting material, into a slot, cut out of the gel after cooling. Separations were carried out at 4°C for nine hours at 10V/cm. Under these conditions, despite the lack of supporting material in the slots, the electrodecantation effect observed by Smithies[21] in horizontal gels was negligible. Upon termination of a run, the gels were sliced in half and stained with Amido Black 10B.

Hemoglobin samples for linkage studies were collected by absorbing a drop of tail blood on a 5 × 9 mm piece of Whatman No. 1 filter paper. The pieces of filter paper were inserted into starch gels and separations carried out under conditions identical to those employed for serum, except that the duration of each run was only three hours.

The variant components were identified as β-globulins by the combined paper and starch gel electrophoresis technique described by Poulik and Smithies.[22] They were identified as transferrins by autoradiography with ^{59}Fe, as described by Giblett *et al.*[22]

Results and Discussion.—Several serum protein differences recognizable either electrophoretically or immunologically occur between the lines tested. One of these involves the iron-binding, β-globulin component, usually called transferrin

or siderophilin. Upon separation of the serum components by starch gel electrophoresis, the CBA line differs in this major β-globulin component from all of the other lines thus far tested, in that a protein more negatively charged at pH 8.5 replaces the component found in the other lines (Fig. 1, *A* and *D*). This more rapidly migrating protein found in the CBA line has been designated Trf-1, its slower homologue, Trf-2. These designations will be used also for the phenotypes of individuals which show these components.

Sera from F_1 progeny of crosses between CBA and B/6 contain both Trf-1 and Trf-2 (Fig. 1, *B*). This phenotype will be referred to as Trf-1·2. The two proteins are present in approximately equal amounts, with the Trf-1 band somewhat

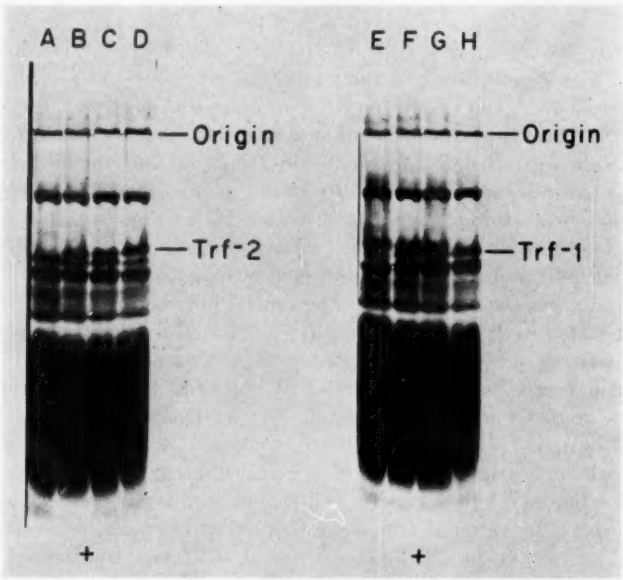


FIG. 1.—Amido Black-stained gels after separation of sera from: (*A*) CBA, (*B*) CBA × B/6 F_1, (*C*) mixture of CBA plus B/6, (*D*) B/6, (*E*) CBA × B/6 F_2 (Trf-2), (*F*) CBA × B/6 F_2 (Trf-1·2), (*G*) CBA × B/6 F_2 (Trf-1·2), (*H*) CBA × B/6 F_2 (Trf-1).

more intense. Each protein seems to be reduced in quantity to about one-half its level in the parental lines, as indicated by comparison of the F_1 sera with a mixture of equal quantities of CBA and B/6 sera (Fig. 1, *C*). Results of tests of F_1 progeny are recorded in Table 1. The results are the same regardless of sex or the direction in which the cross is made, so it may be assumed that the trait is not sex-linked.

Figure 1, *E–H*, shows the electrophoretic patterns of sera from four F_2 progeny. Only the three transferrin phenotypes shown are observed among the F_2 and the backcross progeny; these phenotypes are indistinguishable from those of the parental and F_1 individuals shown in Figure 1, *A*, *B*, and *D*. The simplest hy-

pothesis of genetic control is that the two proteins are determined by a pair of codominant alleles.

The results of the tests of F_2 and backcross progeny are shown in Table 1. Data from males and females have been pooled, tests for homogeneity revealing no significant deviations. The sex ratio of progeny from all crosses is very close to unity. The Chi-square and probability values obtained when the data are tested for fit with the 1:1 and 1:2:1 ratios expected with the codominance hypothesis are also shown in Table 1. The fit is good, although there are excesses of *Trf*-1·2 types in the backcrosses. The data support the hypothesis of genetic control by a pair of codominant alleles. This is further supported by preliminary results from a CBA × DBA/2 cross, in which five F_1's and 42 F_1 × DBA/2 backcrosses show the same phenotypes as the analogous CBA × B/6 crosses.

TABLE 1

DISTRIBUTION OF PHENOTYPES AMONG PROGENY FROM FIRST AND SECOND GENERATION CROSSES BETWEEN CBA AND C57B1/6

Cross		Phenotype					
Dam	Sire	Trf-1	Trf-1·2	Trf-2	Total	χ^2	P
C57B1/6 × CBA		0	26	0	26		
CBA × C57B1/6		0	22	0	22		
Total		0	48	0	48		
F_1 × C57B1/6		0	33	32	65		
C57B1/6 × F_1		0	44	38	82		
Total		0	77	70	147	0.33	0.5
F_1 × F_1		55	115	66	236	1.18	0.5
F_1 × CBA		18	23	0	41		
CBA × F_1		34	43	0	77		
Total		52	66	0	118	1.68	0.2

It is proposed that this locus be designated the transferrin locus (*Trf*), and that the two alleles identified be called Trf^1 and Trf^2, corresponding to the protein types which they control. This symbolism is consistent with that adopted for the hemoglobin locus in mice.[24]

Progeny from backcrosses to B/6 have been analyzed to determine whether the transferrin locus is linked to either of two other loci at which the CBA and B/6 lines differ. The CBA line is agouti, *A*, and has diffuse hemoglobin, Hb^2, the B/6 line is non-agouti, *a*, and has single hemoglobin, Hb^1. The results of these linkage tests are shown in Table 2. The recombination frequencies between the transferrin locus and the hemoglobin and agouti loci are 0.51 ± 0.042 and 0.56 ± 0.041 respectively, neither of which is significantly different from 0.5. In several of the classes in Table 2, rather large deviations from the expected occur. These could indicate viability differences due to interactions between loci or incompatibilities between dam and offspring, or may reflect aberrant segregation of the sort which seems to occur occasionally in certain crosses in mice.[25] Additional backcross progeny are being produced to ascertain whether these deviations have any real basis. It seems probable, however, that the transferrin locus is not in linkage groups I or V, unless it is near or outside *Hk* in I or *Sd* in V.

Figure 2 shows an autoradiogram obtained after electrophoresis of serum samples, containing added ^{59}Fe, from CBA, F_1, F_2 (Trf-1·2), and B/6 individuals, and the same gel stained with Amido Black 10B. Autoradiograms of samples from other F_2 progeny give identical results. There can be little doubt that these proteins

TABLE 2

BACKCROSS LINKAGE RESULTS

Cross			Phenotype			
			Hemoglobin Locus			
			Diffuse		Single	
Dam Sire	Sex		Trf-1·2	Trf-2	Trf-1·2	Trf-2
F_1 × C57B1/6	Male		7	8	11	4
	Female		8	7	7	13
C57B1/6 × F_1	Male		12	16	7	7
	Female		9	3	14	11
Total			36	34	39	35
			Agouti Locus			
			Trf-1·2	Trf-2	Trf-1·2	Trf-2
F_1 × C57B1/6	Male		5	6	13	6
	Female		7	14	8	6
C57B1/6 × F_1	Male		12	11	7	12
	Female		10	9	15	6
Total			34	40	43	30

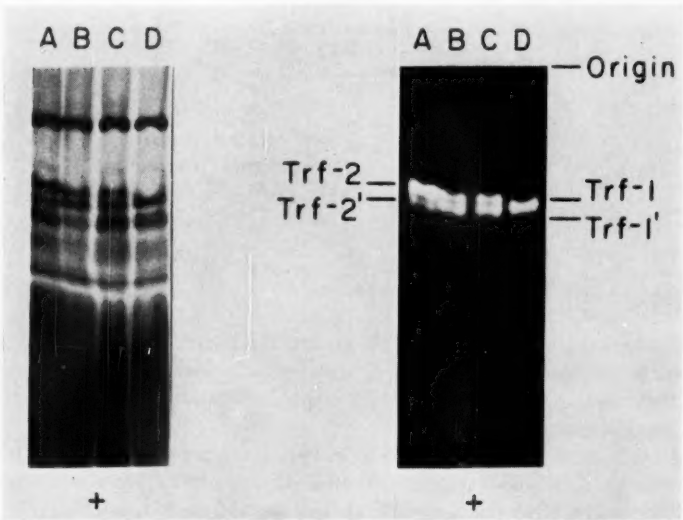


FIG. 2.—Amido Black-stained gel and autoradiogram after separation of samples containing ^{59}Fe. (*A*) B/6, (*B*) CBA × B/6 F_2 (Trf-1·2), (*C*) CBA × B/6 F_1, (*D*) CBA.

are the iron-binding components of the serum, and that only three phenotypes occur in the segregating progeny. It is apparent, however, that in addition to the major components designated Trf-1 and Trf-2 there is associated with each a second, more rapidly migrating, minor component, which also binds ferric ions, and which varies in its mobility as the major component varies. These minor components can be seen in the Amido Black-stained gel shown in Figure 2, as well as in the autoradiogram. Blumberg[26] has also shown a major and a minor transferrin component, but observed no genetic variation in either component. The minor components have been tentatively designated Trf-1′ and Trf-2′. The consistently greater intensity of the Trf-1 band in heterozygotes suggests that this band overlaps a Trf-2′ component. A reduced intensity of Trf-1′ in heterozygotes suggests that less of this component is present in the serum of these animals than in the

Trf-1 phenotype. It appears that the minor components exhibit phenotypic patterns analogous to those of the majors, and that they are controlled by the same kind of genetic mechanism.

The electrophoretic results from the F_2's and backcrosses have been carefully examined for recombination between the major and minor components, e.g., a recombinant type containing Trf-2 and Trf-1′. This might be somewhat difficult to determine conclusively in the Trf-1·2 and Trf-1 phenotypes, but should be readily apparent in the Trf-2 phenotype. No such recombination has been observed in 107 Trf-1, 258 Trf-1·2, and 136 Trf-2 backcross and F_2 progeny. Several explanations are apparent for this phenomenon. The major and minor components could be structurally different and determined by different, but closely linked loci. They could be structurally different, but both controlled by the same locus. They could represent the same structural unit, controlled by the same locus, with one of them conjugated with another unit common to both genotypes, or with major and minor components representing polymeric forms of the same unit. The latter seems the most reasonable explanation, but further physical and chemical analyses, and possibly further genetic studies, will be required to resolve this problem.

The major transferrin components probably correspond to the $\beta_{2\text{-}1}$ component of Clausen and Heremans,[27, 28] identified by immunoelectrophoresis of CBA × DBA/2 F_1's as a transferrin. These workers have reported the identification of only one component having iron-binding properties. Presumably, therefore, the mutation changing the electrophoretic mobility does not affect the antigenic combining site. If this had been altered, two separate transferrin bands should have been seen after immunoelectrophoresis of serum from these mice, which are of the Trf-1·2 phenotype. If the combining site is the same on both Trf-1 and Trf-2, although the components should separate electrophoretically on the agar gels used in immunoelectrophoresis, they would probably not separate far enough to produce any marked change in the precipitin band. The observation of only a single transferrin band further suggests that the minor components might be artifacts due to particular properties of the starch gel or the borate buffer, or again that they have a polymeric relationship to the major components. In immunoelectrophoresis, different polymeric forms of the same unit might precipitate in the same band.

The serum transferrin variants in mice could be useful in several ways. The transferrin locus may be useful as a genetic marker, particularly when its linkage group has been determined. The potential value of detailed chemical analyses of such variants has been discussed. The method of Boettcher *et al.*[29] for isolation of the human transferrins may be applicable for isolation of the mouse transferrins as well, since they have a similar mobility in starch gel and have been shown by Clausen and Heremans[28] to cross-react with antisera against human transferrin. The transferrin variants might be of value in transplantation studies, as markers for detecting repopulation in radiation chimeras, although Popp and Smith[30] were unable to demonstrate repopulation by serum protein-synthesizing cells with rat-into-mouse heterologous chimeras.

Summary.—Sera from mice of 13 inbred lines have been compared by starch gel electrophoresis. The CBA line differs from the other lines tested in the substitution of a more rapidly migrating β-globulin component, a transferrin, for the slower-migrating transferrin component found in the other lines. Sera from 48

F_1 progeny from reciprocal crosses between CBA and C57B1/6 contain both components in approximately equal amounts, but each is reduced in quantity relative to the parental lines. Tests of 501 F_2 and backcross progeny show segregation of the parental and the F_1 transferrin phenotypes with the frequencies expected from genetic control of the trait by a pair of codominant alleles. The locus controlling the trait has been called the transferrin locus (*Trf*), and the alleles, Trf^1, in the CBA line, Trf^2 in the other lines. Tests for linkage to the agouti and hemoglobin loci are negative. Autoradiography with ^{59}Fe has shown that each major transferrin component has associated with it a more rapidly migrating minor transferrin component, which appears to be under control of the same locus.

* This work has been supported in part by U. S. Public Health Service grant No. 2G86-C2.

† The author has received personal support from two National Science Foundation Predoctoral Fellowships.

[1] Benzer, S., in *Chemical Basis of Heredity*, ed. W. D. McElroy and B. Glass (Baltimore: Johns Hopkins Press, 1957), pp. 70–93.

[2] Hunt, J. A., and V. M. Ingram in *CIBA Foundation Symposium: Biochemistry of Human Genetics*, ed. G. E. W. Wolstenholme and C. M. O'Connor (London: J. & A. Churchill Ltd., 1959), pp. 114–143.

[3] Aschaffenberg, R., and R. Drewry, *Nature*, **176**, 218–9 (1955).

[4] Tanford, C., and Y. Nozaki, *J. Biol. Chem.*, **234**, 2874–7 (1959).

[5] Timasheff, S. N., and R. Townend, *J. Am. Chem. Soc.*, **80**, 4433–4 (1958).

[6] Ogston, A. G., and M. P. Tombs, *Biochem. J.*, **66**, 399–403 (1957).

[7] Oudin, J., *Compt. Rend.*, **250**, 770–2 (1960).

[8] Dray, S., and G. O. Young, *Science*, **131**, 738–9 (1960).

[9] Dubiski, S., D. Skalba, A. Dubiska, and A. Kelus, *Nature*, **184**, 1811–2 (1959).

[10] Porter, R. R., *Biochem. J.*, **73**, 119–26 (1959).

[11] Nisonoff, A., F. C. Wissler, and D. L. Woernley, *Federation Proc.*, **19**, 198 (1960).

[12] Smithies, O., *Adv. Prot. Chem.*, **14**, 65–113 (1959).

[13] Hickman, C. G., and O. Smithies, *Genetics*, **43**, 374–85 (1958).

[14] Ashton, G. C., *Nature*, **182**. 945–6 (1958).

[15] *Ibid.*, 1029.

[16] Kristjansson, F. K., *Science*, **131**, 1681 (1960).

[17] Gluecksohn-Waelsch, S., H. M. Ranney, and B. F. Sisken, *J. Clin. Invest.*, **36**, 753–6 (1957).

[18] Thompson, S., J. F. Foster, J. W. Gowen, and O. E. Tauber, *Proc. Soc. Exp. Biol. & Med.*, **87**, 315–7 (1954).

[19] Shreffler, D. C., unpublished results.

[20] Smithes, O., *Biochem. J.*, **61**, 629–41 (1955).

[21] *Ibid.*, **71**, 585–7 (1959).

[22] Poulik, M. D., and O. Smithies, *Biochem. J.*, **68**, 636–43 (1958).

[23] Giblett, E. R., C. G. Hickman, and O. Smithies, *Nature*, **183**, 1589–90 (1959).

[24] Russell, E. S., and P. S. Gerald, *Science*, **128**, 1569–70 (1958).

[25] Michie, D., *J. Genetics*, **53**, 285–294 (1955).

[26] Blumberg, B. S., *Proc. Soc. Exp. Biol. & Med.*, **104**, 25–8 (1960).

[27] Clausen, J., and J. Heremans, *Acta Physiol. Scand.*, **48**, 471–8 (1960).

[28] Clausen, J., and J. Heremans, *J. Immunol.*, **84**, 128–34 (1960).

[29] Boettcher, E. W., P. Kistler, and Hs. Nitschmann, *Nature*, **181**, 490–1 (1958).

[30] Popp, R. A., and L. H. Smith, *J. Nat. Cancer Inst.*, **23**, 395–409 (1959).

A NOTE ON LOCAL PROPERTIES OF SOLUTIONS OF ELLIPTIC DIFFERENTIAL EQUATIONS

By A. P. Calderón and A. Zygmund

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO

Communicated August 23, 1960

We formulate here *pointwise* estimates for solutions of elliptic partial differential equations and systems. The results obtained differ from the familiar ones in that they give information about the behavior of solutions at individual points. More specifically, we obtain two kinds of results. On the one hand, we establish inequalities for solutions and their derivatives at isolated individual points. On the other, we also obtain results of the "almost everywhere" type. Theorems 1 and 2 below summarize the main results.

By $x, y, \ldots$ we denote points $(x_1, x_2, \ldots, x_n)$, $(y_1, y_2, \ldots, y_n)$, $\ldots$ of the n-dimensional Euclidean space E_n, and we write $\|f\|_p$ for $\{\int |f(x)|^p\, dx\}^{1/p}$, where dx stands for the element of volume and the integral is extended over E_n. The symbols $x + y$, λx where λ is a scalar, $|x|$ and $x \cdot y$ have their usual meanings. If $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, where the α_j are nonnegative integers, we will write $|\alpha| = |\alpha_1| + |\alpha_2| + \ldots + |\alpha_n|$, $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$, $\alpha! = \alpha_1!\alpha_2! \ldots \alpha_n!$, $\left(\frac{\partial}{\partial x}\right)^\alpha f = f_\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1}\left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \ldots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}$. The symbol C with various subscripts will stand for constants, not necessarily the same at each occurrence, which depend only (except when otherwise stated) on the variables displayed; dependence on the dimension n however will not be indicated. The functions we consider are complex-valued.

It seems that the notion of differentiability which is most suited for the treatment of problems that concern us is not the classical one. It appears that it is more convenient to estimate the remainder of the Taylor series in the mean with various exponents. This type of differentiability is much more stable than the classical one, in the sense that it is preserved at individual points under various operations such as fractional integration and singular integral transformations.

Let $1 \leq p \leq \infty$, and let u be any number $\geq -n/p$. By $T_u^p(x_0)$ we denote the class of functions $f(x) \epsilon L^p = L^p(E_n)$ such that there exists a polynomial $P(x - x_0)$ of degree strictly less than u (in particular, $P \equiv 0$ if $u \leq 0$) with the property that

$$\{\rho^{-n} \int_{|x - x_0| \leq \rho} |f(x) - P(x - x_0)|^p dx\}^{1/p} \leq A\rho^u (0 < \rho < \infty), \tag{1}$$

with A independent of ρ; if $p = \infty$, the left side here means, as usual, less sup $|f(x) - f(x_0)|$ in $|x - x_0| \leq \rho$. Instead of $T_u^\infty(x_0)$ we write $T_u(x_0)$. If $f \epsilon T_u^p(x_0)$ and there is a polynomial $P(x - x_0)$ of degree less than or equal to u such that the left side of (1) is $o(\rho^u)$ as $\rho \to 0$, we shall say that $f \epsilon t_u^p(x_0)$. Any $f \epsilon L^p$ belongs to $T_u^p(x_0)$ for $u = -n/p$ and any x_0. The familiar result about the Lebesgue set of a function can be expressed by saying that if $f \epsilon L^p$, $p < \infty$, then $f \epsilon t_0^p(x_0)$ for almost all x_0.

The class $T_u^p(x_0)$ is a linear space in which we introduce a norm. The norm of an $f \epsilon T_u^p(x_0)$ is denoted by $T_u^p(x_0, f)$ and is defined as the sum of $\|f\|_p$, the moduli of

the coefficients of the polynomial $P(x - x_0)$, and the least admissible value of A in (1).

Let Q be a closed set. We shall say that the bounded function f belongs to the class $B_u(Q)$, $u > 0$, if there exist bounded functions f_α, $|\alpha| < u$, such that

$$f_\alpha(x + h) = \sum_{|\beta| < u - |\alpha|} \frac{h^\beta}{\beta!} f_{\alpha+\beta}(x) + R_\alpha(x, h)$$

for all x and $x + h$ in Q, with $|R_\alpha(x, h)| \leq C |h|^{u-|\alpha|}$. We shall say that f belongs to $b_u(Q)$, if there exist functions f_α with $|\alpha| \leq u$ such that

$$f_\alpha(x + h) = \sum_{|\beta| \leq v - |\alpha|} \frac{h^\beta}{\beta!} f_{\alpha+\beta}(x) + R_\alpha(x, h)$$

for all x and $x + h$ in Q, with $|R_\alpha(x, h)| \leq C |h|^{u-|\alpha|}$, and, in addition, $R_\alpha(x,h) = o(|h|^{u-|\alpha|})$ as $h \to 0$, uniformly in $x \in Q$.

Given a nonnegative integer k, we denote by $L_k{}^p$ the class of functions in L^p with distribution derivatives of order $\leq k$ in L^p. The norm $\|f\|_{p,k}$ of an $f \in L_k{}^p$ is defined as the sum of the norms in L^p of f and of its derivatives of order $\leq k$.

Let $f = (f_1, f_2, \ldots, f_r)$ be a vector valued function and

$$\mathfrak{L}f = \sum_{|\alpha| \leq m} a_\alpha(x) \left(\frac{\partial}{\partial x}\right)^\alpha f = g,$$

where $g = (g_1, g_2, \ldots, g_s)$ is a vector valued function of s components, $s \geqslant r$, and $a_\alpha(x)$ are $s \times r$ matrices. We say that the operator $\mathfrak{L}$, or the equation $\mathfrak{L}f = g$, is *elliptic at the point* x_0 if the characteristic matrix

$$\sum_{|\alpha| = m} a_\alpha(x)\xi^\alpha \tag{2}$$

has the property that, for $|\xi| = 1$,

$$\det\left[\left(\sum_{|\alpha| = m} a_\alpha{}^*(x_0)\xi^\alpha\right)\left(\sum_{|\alpha| = m} a_\alpha(x_0)\xi^\alpha\right)\right] \geqslant \mu(x_0) > 0,$$

where a^* denotes the conjugate transpose of a. We call $\mu(x_0)$ an *ellipticity constant* of $\mathfrak{L}$ at x_0.

THEOREM 1. *Let $\mathfrak{L}f = g$ be an equation of order m with coefficients in $T_u(x_0)$, $u > 0$, which is elliptic at x_0 in the sense just defined. Let $1 < p < \infty$, $u \geqslant v \geqslant -n/p$, v nonintegral. If $f \in L_m{}^p$, $g \in T_u{}^p(x_0)$, in the sense that their components belong to these spaces, then*

$$\text{(i) } T^q_{v+m-|\alpha|}\left(x_0, \left(\frac{\partial}{\partial x}\right)^\alpha f_j\right) \leq C\left[\sum_{i=1}^{s} T_v{}^p(x_0, g_i) + \sum_{i=1}^{r} \|f_i\|_{p,m}\right] \tag{3}$$

for $j = 1, 2, \ldots, r$ and $|\alpha| \leqslant m$, where $\frac{1}{p} \geqslant \frac{1}{q} \geqslant \frac{1}{p} - \frac{m - |\alpha|}{n}$ if $\frac{1}{p} - \frac{m - |\alpha|}{n} > 0$, $p \leq q \leq \infty$ if $\frac{1}{p} < \frac{m - |\alpha|}{n}$, and $p \leq q < \infty$, if $\frac{1}{p} = \frac{m - |\alpha|}{n}$, and C depends only on v, p, r, s, $\mu(x_0)$ and the least upper bound of the norms in $T_u(x_0)$ of the coefficients of $\mathfrak{L}$.

(ii) *If, in addition, the leading coefficients of the equation are uniformly continuous and the equation is uniformly elliptic in the sense that the constant of ellipticity $\mu(x)$ is*

bounded away from 0, *then the quantity* $\sum \|f_i\|_{p,m}$ *on the right of* (3) *can be replaced by* $C\left[\sum_{i=1}^{r} \|f_i\|_p + \sum_{i=1}^{s} \|g_i\|_p\right]$, *where* C *depends on* $\mathfrak{L}$ *only.*

(iii) *If* $g \in t_v^p(x_0)$, *then* $\left(\frac{\partial}{\partial x}\right)^\alpha f$ *belongs to* $t^q_{v+m-|\alpha|}(x_0)$ *with the same* q *as in part* (i).

THEOREM 2. *Let* $\mathfrak{L}f = g$ *be an equation which is elliptic at all points belonging to a set* Q *of positive measure, and whose coefficients belong to* $T_u(x_0)$, $u \geq 1$, *for all* x_0 *in* Q. *Let* v *be a positive integer not larger than* u. *Then, if* $g \in T_v^p(x_0)$, $1 < p < \infty$, *for all* $x_0 \in Q$, *and* $f \in L_m^p$, *the functions* $\left(\frac{\partial}{\partial x}\right)^\alpha f$, $|\alpha| \leq m$, *belong to* $t^q_{v+m-|\alpha|}(x_0)$ *for almost all* x_0 *in* Q, *where* q *is the same as in part* (i) *of Theorem 1.*

THEOREM 3. *Let* $\mathfrak{L}f = g$ *be an equation of order* m *with coefficients in* $T_u(x_0)$, $u > 0$, *for all* x_0 *in a closed set* Q. *Let the norms of the coefficients in* $T_u(x_0)$ *be bounded in* Q *and* $\mathfrak{L}$ *be uniformly elliptic in* Q. *If* $g \in T_v^p(x_0)$ *for all* x_0 *in* Q *and* $\sum_{i=1}^{s} T_v^p(x_0, g_i)$ *is bounded in* Q, $1 < p < \infty$, v *is positive and nonintegral,* $-m < v \leq u$ *and* $f \in L_m^p$, *then the function* f *belongs to* $B_{v+m}(Q)$. *If, in addition,* $g \in t_v^p(x_0)$ *for all* $x_0 \in Q$, *then* $f \in b_{v+m}(Q)$.

It has already been observed[1] that, roughly speaking, a differential operator is the composition of fractional differentiation and a singular integral transformation. Correspondingly, the method of proof of Theorems 1, 2, 3 consists in deriving estimates for solutions of differential equations from estimates for fractional integrals and singular integral transforms. These estimates seem to be of independent interest and we state some of them here.

Let f be a tempered distribution in E_n. The fractional integral of order u of f will be denoted by $J^u f$ and defined by[2]

$$\widehat{J^u f} = (1 + 4\pi^2 |x|^2)^{-u/2} \hat{f},$$

where $\hat{f}$ stands for the Fourier transform, $\hat{f}(x) = \int_{E_n} e^{-2\pi i(x \cdot y)} f(y) dy$, of f.

THEOREM 4. *Let* $u \geq -n/p$, $v > 0$, $u + v$ *nonintegral,* $p > 1$. *Then* J^v *maps continuously*

(i) $T_u^p(x_0)$ *into* $T^q{}_{v+u}(x_0)$ *provided,*

a) $\frac{1}{p} \geq \frac{1}{q} \geq \frac{1}{p} - \frac{v}{n}$, *if* $p < \frac{n}{v}$, b) $p \leq q \leq \infty$, *if* $\frac{n}{v} < p \leq \infty$, c) $p \leq q < \infty$ *if* $\frac{n}{v} = p$; (ii) $t_u^p(x_0)$ into $t_u^q{}_{+v}(x_0)$, *with* u *and* v *as in* (i).

THEOREM 5. *Let* u *and* v *be nonnegative integers. Then, if* $f \in T_u^p(x_0)$, $1 < p < \infty$, *for all* x_0 *in a set* Q *of positive measure, we have* $J^v f \in t_u^q{}_{+v}(x_0)$ *for almost all* x_0 *in* Q, p *and* q *being related as in part* (i) *of Theorem 4. If* $u + v \geq 1$, *the assertion is valid for* $1 < p \leq \infty$.

This theorem asserts, in particular, that if u is a positive integer and $1 < p \leq \infty$ the condition $f \in T_u^p(x_0)$ implies $f \in t_u^p(x_0)$ almost everywhere (the case $u = 1$ and $p = \infty$ is the familiar theorem of Rademacher-Stepanov). Related to it are the following two theorems.

THEOREM 6. *Let* f *have first order derivatives* f_j *in* L^p, *and let* $f_j \in T_u^p(x_0)$, $j = 1, 2,$

$\ldots, n, u \geqslant -n/p, u \neq -1$. *If* (1) $1 \leq p < n$ *and* $f \in L^q$, *with* $\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$, *then* $f \in T_u{}^q{}_{+1}(x_0)$ *and* $T_u{}^q{}_{+1}(x_0, f) \leq C \sum_{j=1}^{n} T_u{}^p(x_0, f_j)$; (2) *if* $n < p \leq \infty$ *and* $f \in B(E_n)$, *then* $f \in T_{u+1}(x_0)$ *and* $T_{u+1}(x_0, f) \leq C_{p,u}\,[B(f) + \sum_{j=1}^{n} T_u{}^p(x_0, f_j)]$; (3) *If* $f \in L^r$, *with* $\frac{1}{p} \geqslant \frac{1}{r} > \frac{1}{p} - \frac{1}{n}$, *then* $f \in T_u{}^r{}_{+1}(x_0)$ *and* $T_u{}^r{}_{+1}(x_0, f) \leq C_{r,u}[\|f\|_r + \sum_{j=1}^{n} T_u{}^p(x_0, f_j)]$. (4) *The preceding statements hold with the spaces* T *replaced by the corresponding spaces* t.

THEOREM 7. *If* $f \in L_k{}^p$, $1 \leq p < \infty$, $k = 0, 1, 2, \ldots$, *then* $f \in t_k{}^q(x_0)$ *for almost all* x_0, *with* $\frac{1}{p} \geqslant \frac{1}{q} \geqslant \frac{1}{p} - \frac{k}{n}$ *if* $p < n/k$; $p \leq q \leq \infty$ *if* $p > n/k$; *and* $p \leq q < \infty$ *if* $p = n/k$. *Moreover, given* $\epsilon > 0$ *there is a function* $g(x)$ *with continuous derivatives of orders* $\leq k$, *such that* $f(x) = g(x)$ *outside a set of measure* $< \epsilon$.

We shall now consider singular integral operators of the following form:

$$\mathcal{K} f = a(x)f(x) + \int k(x, x - y)f(y)dy, \tag{4}$$

where $a(x)$ is a bounded measurable function and $k(x, z)$ is homogeneous of degree $-n$ with respect to z (i.e., $k(x, \lambda z) = \lambda^{-n} k(x, z)$ for $\lambda > 0$), and further $k(x, z)$ has for each x mean value 0 on $|z| = 1$. In addition we assume that $k(x, z)$ is infinitely differentiable with respect to z and is uniformly bounded on $|z| = 1$. The integral in (4) must be taken in the principal-value sense. Associated with the operator $\mathcal{K}$ is its *symbol* $\sigma(\mathcal{K})$ which is defined by the equation

$$\sigma(\mathcal{K}) = a(x) + k(x, \hat{z}),$$

where $k(x, \hat{z})$ is the Fourier transform of $k(x, z)$ with respect to z.

An operator $\mathcal{K}$ as above is said to belong to the class $T_u(x_0)$, $u \geqslant 0$, if $\sigma(\mathcal{K})$ and its derivatives with respect to the coordinates of z of orders $\leq 2n + u$ belong to $T_u(x_0)$ for each $z \neq 0$, uniformly on $|z| = 1$. The norm $T_u(x_0, \mathcal{K})$ of an operator of the class $T_u(x_0)$ is, by definition, the least upper bound of the norms in $T_u(x_0)$ of $\sigma(\mathcal{K})$ and its derivatives with respect to z of orders less than or equal to $2n + u$ evaluated on $|z| = 1$. Operators of class $t_u(x_0)$ are defined analogously.

THEOREM 8. $\mathcal{K}$ *be a singular integral operator of the class* $T_u(x_0)$, $1 < p < \infty$, $v \geqslant -n/p$, $v \neq 0, 1, 2, \ldots$, $u \geqslant v$. *Then* $\mathcal{K}$ *maps continuously* $T_v{}^p(x_0)$ *into* $T_v{}^p(x_0)$, *with norm less than or equal to* $C_{p,u}T_u(x_0, \mathcal{K})$. *The corresponding results for operators of class* $t_u(x_0)$ *and the spaces* $t_v{}^p(x_0)$ *is also valid.*

THEOREM 9. *Let* u *be a nonnegative integer and* f *a function belonging to* $T_u{}^p(x_0)$, $1 < p < \infty$, *for all* x_0 *in a set* Q *of positive measure. Then there exists a subset* $\bar{Q}$ *of* Q *such that* $Q - \bar{Q}$ *has measure* 0 *and such that for every singular integral operator* $\mathcal{K}$ *belonging to* $T_u(x_0)$, $x_0 \in \bar{Q}$, $\mathcal{K}f$ *belongs to* $T_u{}^p(x_0)$.

We conclude with an additional result about the solutions of the equation $\mathfrak{L}f = g$. It is well known that Theorem 1 is not true when u is a nonnegative integer, and the simplest illustration is the equation $\Delta f = g$: if g is merely continuous, f is not necessarily twice differentiable. The only thing we can say then is that f has

continuous first order derivatives $f_j = \frac{\partial}{\partial x_j} f$, $j = 1, 2, \ldots n$, which, in turn, satisfy a Hölder condition of any order less than 1 or, more generally, have a modulus of continuity $\omega(|h|) = \left(|h| \log \frac{1}{|h|}\right)$.

It is, however, not difficult to show, and the result is a special case of Theorem 10 below, that if g is continuous at a given point x_0, then any solution of $\Delta f = g$ has continuous derivatives f_j near x_0 and the f_j satisfy at that point the condition of *smoothness:*

$$f_j(x_0 + h) + f_j(x_0 - h) - 2 f_j(x_0) = o(|h|),\ h \to 0, \tag{5}$$

a condition first introduced by Riemann in the theory of trigonometric series.

This condition can be interpreted as the differentiability at $h = 0$ of the 'even part' $\frac{1}{2}[f_j(x_0 + h) + f_j(x_0 - h)]$ of $f_j(x_0 + h)$, and from this it is easy to deduce that the 'odd part' $\frac{1}{2}[f(x_0 + h) - f(x_0 - h)]$ of f has a second differential at $h = 0$. This fact gives a clue to the situation in the general case. We may add, and the result is familiar enough, that if a continuous function $g(x)$ satisfies the condition $g(x + h) + g(x - h) - 2g(x) = o(|h|)$ uniformly in a domain, then in every compact subdomain g has a modulus of continuity $o\left(|h| \log \frac{1}{|h|}\right)$, so that the result about the modulus of continuity of the derivatives f_j in the case of the equation $\Delta f = g$ is a corollary of (5).

We shall say that $f(x)$ belongs to the class $\Lambda_u^p(x_0)$, where u is a nonnegative integer, if $\frac{1}{2}[f(x_0 + h) - (-1)^u f(x_0 - h)]$ belongs to $T_u^p(0)$ as a function of h. We shall say that $f(x)$ belongs to $\mathbf{M}_u^p(x_0)$, if $\frac{1}{2}[f(x_0 + h) + (-1)^u f(x_0 - h)]$ belongs to $T_u^p(0)$. On replacing here $T_u^p(0)$ by $t_u^p(0)$ we obtain the definitions of $\lambda_u^p(x_0)$ and $\mu_u^p(x_0)$.

THEOREM 10. *Let $\mathfrak{L}f = g$ be an equation of order m with coefficients in $T_v(x_0)$ which is elliptic at x_0. Let u be a nonnegative integer, $v > u$. Then, if m is even and $g \in T_w^p(x_0)$, $1 < p < \infty$, $w > u - 1$, and $g \in \Lambda_u^p(x_0)$, the function $f_\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha f$ belongs $\Lambda_{u+m-|\alpha|}^q$, where p and q are related as in Theorem 1. If m is odd, the same conclusions hold with the classes $\Lambda_u^p(x_0)$ replaced by $\mathbf{M}_u^p(x_0)$. Analogous conclusions hold with classes Λ replaced by λ, and $\mathbf{M}$ by μ, as the case may be.*

The proof gives also inequalities for the norms.

[1] Calderón, A. P., and A. Zygmund, "Singular Integral Operators and Differential Equations," *Amer. J. of Math.*, **79**, 901–921 (1957).

[2] For further properties of the operator J^u see N. Aronsajn and K. T. Smith, "Theory of Bessel Potentials," in *Studies of Eigenvalue Problems*, (University of Kansas, Department of Mathematics, 1959), Technical Report 22, part I, pp. 1–113, or A. P. Calderón, "Singular Integrals," course notes of lectures given at the M.I.T., 1958–1959.

*DEGREE OF APPROXIMATION BY BOUNDED HARMONIC FUNCTIONS**

By J. L. Walsh

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated August 25, 1960

The writer has recently published several notes on degree of approximation on a Jordan curve[1, 2] or arc[3] by polynomials in the complex variable, and also a paper on approximation by bounded harmonic functions,[4] all with emphasis on properties invariant under conformal transformation. The last-named paper admits of extensions (i) to more general degrees of approximation and (ii) to approximation on a Jordan arc; the object of the present note is briefly to set forth these extensions.

Theorem 1. *Let $\epsilon_1, \epsilon_2, \ldots$ be a sequence of positive numbers approaching zero, where we suppose $\epsilon_{[n/\lambda]} = O(\epsilon_n)$ for every positive integral λ (here $[m]$ denotes the largest integer not greater than m) and where for every r, $0 < r < 1$, we have $r^n = O(\epsilon_n)$. Let E be an analytic Jordan curve in the z-plane containing the origin in its interior, and let D be a region containing E. Let the function $u(z)$ be defined on E, let the functions $u_n(z)$ be harmonic in D, and let $(n = 1, 2, 3, \ldots)$*

$$|u(z) - u_n(z)| \leqq A_1\epsilon_n, \qquad z \text{ on } E, \tag{1}$$

$$|u_n(z)| \leqq A_2R^n, \qquad z \text{ in } D, \tag{2}$$

be satisfied. Then there exist harmonic polynomials $p_n(z, 1/z)$ in z and $1/z$ of respective degrees n such that

$$|u(z) - p_n(z, 1/z)| \leqq A_3\epsilon_n, \; z \text{ on } E. \tag{3}$$

Here and below the letter A with subscript denotes a positive constant independent of n and z, a constant which may change from one usage to another.

Map one-to-one and conformally E onto the unit circumference E_1: $|w| = 1$ in the w-plane; then the map is one-to-one and conformal also in regions D_1 (a subregion of D) and D_2 containing E and E_1 respectively. As in the proof of Theorem 2.2 of reference 4, and by use of Lemma 1.2 of reference 4, there exist functions $\phi_n(w)$ *single-valued* and analytic in D_2 whose transforms $\Phi_n(z) \equiv \Phi_{n1}(z) + i\Phi_{n2}(z)$ satisfy the conditions $\Phi_{n1}(z) \equiv u_n(z)$ on E,

$$|\Phi_n(z)| \leqq A_4R_1^n, \qquad z \text{ in } D_3, \tag{4}$$

$$|u(z) - \Phi_{n1}(z)| \leqq A_1\epsilon_n, \qquad z \text{ on } E, \tag{5}$$

where D_3 is an annular subregion of D_1 containing E. In a suitably chosen annular subregion D_4 of D_3 containing E the components (defined in terms of the Cauchy integral) of the respective $\Phi_n(z)$ can be approximated by polynomials in z and polynomials in $1/z$, as in reference 2, in such a way that we have

$$|\Phi_n(z) - P_{nk}(z, 1/z)| \leqq A_5R_1^n\sigma_1^k, \qquad 0 < \sigma_1 < 1, \; z \text{ on } E, \tag{6}$$

where $P_{nk}(z, 1/z)$ is a polynomial in z and $1/z$ of degree k, and σ_1 is independent of k, n, and z. Here we choose the integer λ with $R_1\sigma_1^{\lambda} < 1$; it then follows from equations (5) and (6) that the real parts $p_{n,\lambda n}(z, 1/z)$ of the polynomials $P_{n,\lambda n}(z, 1/z)$ satisfy

$$|u(z) - p_{n,k}(z, 1/z)| \leqq A_6\epsilon_k, \qquad z \text{ on } E, \tag{7}$$

for the sequence $k = n, 2n, 3n, \ldots$. The $p_k(z, 1/z)$ may now be defined for every k by the equations

$$p_k(z, 1/z) \equiv p_{n,\lambda n}(z, 1/z), \qquad \lambda n \leqq k < (\lambda + 1)n,$$

$$p_k(z, 1/z) \equiv 0, \qquad 1 \leqq k < \lambda,$$

and equation (3) follows from (7) by the hypothesis on the ϵ_n.

A somewhat simpler result than Theorem 1 deserves to be stated:

COROLLARY. *Under the conditions of Theorem 1, let D contain both E and the interior of E. Then the polynomials $p_n(z, 1/z)$ may be chosen as harmonic polynomials in z so that equation (3) is valid.*

Under these conditions we map E and its interior onto E_1 and its interior; Jordan regions D_1, D_3, and D_4 may be chosen to contain E and its interior, and the $P_{nk}(z, 1/z)$ satisfying (6) may be chosen as polynomials in z alone, whose real parts $p_n(z, 1/z)$ are harmonic polynomials in z and satisfy equation (7). Then equation (3) follows as in the proof of Theorem 1.

The Corollary is in fact valid (compare reference 2) if D is an arbitrary Jordan region and if E is an arbitrary closed point set interior to D; we use Lemma 1.2 of reference 4, and conformal transformation onto the w-plane is not necessary in the proof.

The Corollary admits a converse:

THEOREM 2. *Let the ϵ_n satisfy the conditions of Theorem 1, let E be a Jordan curve, and let D be a bounded region containing E and its interior. If $u(z)$ is defined on E, and if harmonic polynomials $p_n(z)$ in z of respective degrees n exist satisfying*

$$|u(z) - p_n(z)| \leqq A_1\epsilon_n, \qquad z \text{ on } E, \tag{8}$$

then for suitably chosen R we have

$$|p_n(z)| \leqq A_2R^n, \qquad z \text{ in } D. \tag{9}$$

Thus the $p_n(z)$ satisfy the conditions prescribed for the $u_n(z)$ in the hypothesis of the Corollary.

It follows from (8) that the $p_n(z)$ are uniformly bounded on and interior to E. Let Γ be a circle interior to E, and let Γ_1 be a concentric circle whose radius is R times as large. If R is suitably chosen, D lies in Γ_1. The conclusion (9) now follows by Lemma 1.3 of reference 4, or by a more specific lemma due to Szegö (*loc. cit.*).

The hypothesis of Theorem 1 is invariant under one-to-one conformal (i.e. analytic) transformation of E, which implies also one-to-one conformal transformation of some region D containing E. The hypothesis remains invariant also if (as in the Corollary to Theorem 1 and in Theorem 2) the region D is required to contain both E and its interior. Under the latter condition the class of functions that can be approximated on E (not necessarily analytic) by harmonic polynomials in z of respective degrees n with degree of approximation $O(\epsilon_n)$ is also invariant under one-to-one conformal transformation of E and its interior; if E is an analytic Jordan curve, this class on E itself is thus the set of transforms of functions (necessarily continuous, so the Dirichlet problem for $|z| < 1$ has a solution) which can

be approximated on $|z| = 1$ by trigonometric polynomials in $\theta = -i \log z$ of respective degrees n with degree of approximation $O(\epsilon_n)$.

A further remark is in order:

COROLLARY. *Under the conditions of Theorem 1, the polynomials $p_n(z, 1/z)$ can be chosen as harmonic polynomials in z alone of respective degrees n.*

As in the proof of Theorem 1, map one-to-one and conformally the curve E and its interior onto E_1: $|w| = 1$ and its interior, say $w = \phi(z)$, $z = \psi(w)$. By the invariance of the hypothesis of Theorem 1 and by Theorem 1 itself, there exist harmonic polynomials $p_n(w, 1/w)$ in w and $1/w$ of respective degrees n such that

$$|u[\psi(w)] - p_n(w, 1/w)| \leqq A_3\epsilon_n, \qquad w \text{ on } E_1.$$

However, on E_1: $|w| \equiv \rho = 1$ a harmonic polynomial in w and $1/w$ of degree n is also a harmonic polynomial in w alone of degree n; for every k we have on E_1

$$\rho^k(a_{1k} \cos k\theta + b_{1k} \sin k\theta) + \rho^{-k}(a_{2k} \cos k\theta + b_{2k} \sin k\theta) \equiv \rho^k[(a_{1k} + a_{2k}) \cos k\theta + (b_{1k} + b_{2k}) \sin k\theta].$$

By the invariance of the property expressed by equation (8), the present corollary follows.

The proof just given shows also that *if E is the unit circumference, and if equation (3) is satisfied, then equation (8) is also satisfied if the $p_n(z)$ are suitably chosen.* The corresponding question is still open if E is an arbitrary analytic Jordan curve.

We turn now from approximation on Jordan curves and regions to approximation on Jordan arcs, as in reference 3.

THEOREM 3. *Let the ϵ_n satisfy the conditions of Theorem 1, let E be an analytic Jordan arc, and let D be a region containing E. Let the function $u(z)$ be defined on E, let the functions $u_n(z)$ be harmonic in D, and let (1) and (2) be satisfied. Let $z = \phi(w)$ map E one-to-one and conformally onto E_o: $-1 \leqq w \leqq 1$. Then the function $u[\phi(\cos\theta)]$ can be approximated by trigonometric polynomials in θ of degree n with degree of approximation $O(\epsilon_n)$.*

Conversely, if $u(z)$ is given on E and $u[\phi(\cos\theta)]$ can be approximated by trigonometric polynomials in θ of degree n with degree of approximation $O(\epsilon_n)$, then functions $u_n(z)$ harmonic in D exist satisfying (1) and (2); the $u_n(z)$ may be chosen as harmonic polynomials in z of respective degrees n.

Since the original conditions on the $u_n(z)$ are invariant under one-to-one conformal transformation of D, or of a subregion of D containing E, the method of proof of the Corollary to Theorem 1 shows that on E_o the function $u[\phi(w)]$ can be approximated by harmonic polynomials in w of degree n with degree of approximation $O(\epsilon_n)$. On E_o these harmonic polynomials of degree n are also polynomials of degree n in the real variable w, and the first part of Theorem 3 follows[3] by use of the classical transformation $w = \cos\theta$ as in the Bernstein-Jackson-Montel-de la Vallée-Poussin theory of trigonometric approximation.

To prove the converse, we note that $u[\phi(w)]$ can be approximated on E_o by polynomials of degree n in the real variable w with degree of approximation $O(\epsilon_n)$, again by the classical theory of trigonometric approximation. On E_o these polynomials are also polynomials in the complex variable w of degree n, and by the Bernstein lemma (§3 of reference 4) as proved by S. Bernstein these polynomials satisfy the analogue of equation (2) in an arbitrary bounded region D_1 containing

E_ρ. Thus the real parts of these polynomials in w are harmonic polynomials in w of respective degrees n which satisfy in the w-plane the analogues of equation (1) and equation (2). By the invariance under conformal transformation of the property expressed by equations (1) and (2) we deduce equations (1) and (2) in the main conclusion of the second part of Theorem 3, except that now D is the image in the z-plane of a suitably chosen region in the w-plane rather than an arbitrary region of the z-plane containing E. However, by the methods of proof of the Corollary to Theorem 1 and of Theorem 2 it follows that in equation (1) the $u_n(z)$ may be chosen as harmonic polynomials in z of respective degrees n, and equation (2) is valid in an arbitrary bounded region.

THEOREM 4. *Under the conditions and with the notation of Theorem 3 on the ϵ_n, D, and E, let s be arc-length measured along E algebraically from the mid-point of E, and let $2l$ be the total length of E. Then $f(z)$ can be approximated on E by functions $u_n(z)$ harmonic in D which satisfy (1) and (2) when and only when the function $u(z) \equiv u_2(l \cos \theta_1) \equiv u_2(s)$ can be approximated by trigonometric polynomials in θ_1 with degree of approximation $O(\epsilon_n)$.*

Theorem 4 follows from Theorem 3 of reference 3, which shows essentially that trigonometric approximation in θ of degree $O(\epsilon_n)$ implies trigonometric approximation in θ_1 of degree $O(\epsilon_n)$, and conversely.

It is proved in Theorem 2 that if the $u_n(z)$ are harmonic polynomials in z of respective degrees n, inequality (2) is a consequence of inequality (1), provided E is a Jordan curve. This conclusion is false if E is an analytic Jordan arc, as is shown by the counter-example E: $-1 \leqq z \leqq 1$, $u(z) \equiv 0$, $u_n(z) \equiv n^n xy$ for all n. The $u_n(z)$ are not uniquely determined by their values on E. Likewise if the $u_n(z)$ are harmonic polynomials in z and $1/z$ of respective degrees n and if E is E_1: $|z| = 1$, it is not true that (1) implies (2); a counter-example is $u_n(z) \equiv n^n (\rho^n \cos n\theta - \rho^{-n} \cos n\theta)$, where ρ and θ are the usual polar coordinates.

However, a harmonic polynomial $p_n(z, 1/z)$ in z and $1/z$ of degree n is not uniquely determined by its values on E_1. If a sequence $p_n(z, 1/z)$ of such (real) polynomials is bounded on E_1, each polynomial is a trigonometric polynomial on E_1, and can be expressed on E_1 as a polynomial $P_n(z, 1/z)$ in z and $1/z$ of degree n; the latter polynomials satisfy $|P_n(z, 1/z)| \leqq A_1 R^n$, $R > 1$, in an arbitrary annulus $1/R \leqq |z| \leqq R$, and the real parts of the $P_n(z, 1/z)$ are harmonic polynomials in z and $1/z$ of degree n that satisfy this same inequality and coincide with $p_n(z, 1/z)$ on E_1.

Similarly, in the particular case that E is $-1 \leqq z \leqq 1$, a (real) harmonic polynomial $p_n(z)$ in z of degree n is on E also a polynomial in x (where $z = x + iy$) of degree n on E, thus a polynomial $P_n(z)$ in the complex variable z of degree n. When so interpreted, it follows from the original form of Bernstein's lemma as proved by S. Bernstein that $|P_n(z)| \leqq A_1$, z on E, implies $|P_n(z)| \leqq A_1 R^n$, for z on and interior to the ellipse $|z^2 - 1| = (R + 1/R)/2$. The real part of $P_n(z)$ is a harmonic polynomial $q_n(z)$ of degree n which coincides with $p_n(z)$ on E and which satisfies $|q_n(z)| \leqq A_1 R^n$ on and interior to that ellipse.

* This research was supported in part by the United States Air Force Office of Scientific Research, Air Research, and Development Command.

[1] Walsh, J. L., *Math. Zeit.*, **72**, 47–52 (1959).

[2] Walsh, J. L., these PROCEEDINGS, **45**, 1528–1533 (1959).

[3] *Ibid.*, **46**, 981–983 (1960).

[4] Walsh, J. L., *Jour. de Math. pures et appl.*, **39**, (1960).

*CRITIQUE OF CRITICAL TABLES**

By P. W. Bridgman

DEPARTMENT OF PHYSICS, HARVARD UNIVERSITY

Communicated August 17, 1960

The need for critical tables has grown out of the spectacular fecundity of modern experimental research. Some fifty thousand scientific periodicals throughout the world present a wealth of material with which it is impossible for any one individual to come into mere physical contact, much less to digest and evaluate. A large part of the material thus published which is pertinent to critical tables comprises numerical data of various sorts, perhaps for the most part values of the various physical parameters which describe the properties of different materials, but also numerical values for fundamental constants such as the velocity of light or the intensity of gravity. This numerical material is subject to extensive duplication, for many measurements are made independently of each other. Furthermore, techniques are continually being improved so that continually better values are being obtained for ostensibly the same parameter. It is the primary function of a set of critical tables to subject this welter of numerical material to some sort of critical evaluation, which will enable the interested user to form a judgment as to the "best" value for any parameter, and also to make some estimate of the "probable" error of the value selected as the "best."

It is the purpose of this introductory talk to inquire into some of the principles which should control any attempt to select a "best" value or to estimate how good the "best" value may be expected to be. This inquiry will be both general and also particular, in so far as special kinds of parameters demand special treatment. I cannot hope to give you anything philosophically profound, or to do more than to present considerations obvious to every one of you who is willing to give the matter a little thought.

It is, I think, usually implied that not only should any critical tables endeavor to present the "best" values, but that they should also endeavor to find, within limits, the "correct" values. That is, the usual implication is that there is an "objective" background to all our measurement, and that the ultimate purpose of our measurements is to find the "objective" values of the various parameters, such as the velocity of light, which are true and correct independent of the measuring instrument or of the measuring agent. Now it may well be that something of this sort is an ideal, and a proper ideal of our measurements, but I think that examination of what we actually do when making measurements and determining the numerical values of our parameters will show that any such ideal is at best a limiting ideal which we never actually attain in practice and in fact never can attain. For it is obvious that the making of numerical measurements and the assessing of their probable excellence is unavoidably a human enterprise, and that all we can attain at any epoch is a consensus of the best opinion of qualified experts at that epoch. Now, with all due deference to my audience, the opinion of qualified experts is subject to the hazards of human frailty, and our evaluation of what can be accomplished by critical tables must take into account as one of the factors this element of human frailty. The importance of this factor may be compara-

tively minor in a field which has long been cultivated and in which we feel confident of the theoretical foundations, but in a new field in which the foundations are not well laid, human psychology may play a surprisingly large role. Rutherford used to enjoy telling how the early measurements of e, the charge of the electron, clustered around the value 3×10^{-10}, until he made the psychological break-through by publishing the value 4.8×10^{-10}, around which all subsequent values have clustered.

The same thing is true, although to a much less degree, with regard to other fundamental physical constants. A study of the "best" values of these constants which have been published from time to time is informative. Thus the value published in *International Critical Tables* in 1926 for Planck's constant, h, was 6.554×10^{-27}, with an indicated uncertainty of 0.001. Revised critical values published in 1929 and 1933 were 6.547 and 6.542×10^{-27}, differing from the earlier value by considerably more than the indicated uncertainty. In 1941 came the psychological break-through by Birge, who raised the "best" value to 6.624, around which subsequent estimates have more or less clustered. The maximum estimated probable error of the three earlier values was much smaller than the jump in the absolute value between 1933 and 1940. Something similar has happened to the velocity of light; the estimated uncertainty in the I.C.T. value of 1926 was only half the shift in accepted absolute value since then.

Various more or less automatic and impersonal methods have been proposed for evaluating the probable error in a numerical value obtained by adjustment from various independent sources, but these have obviously not been successful up to the present, and there seems no completely "objective" method of estimating even the *limits* of error of any specific numerical parameter, nor can any precise logical meaning be given to the "probability" which we refer to when we talk about "probable" error. It seems impossible here to avoid the play of strong personal factors. It is perhaps more common for the individual investigator to overestimate the accuracy of his own measurements, but on the other hand, there are individuals who, recognizing this tendency, react by underestimating their own accuracy. Nevertheless, in spite of all this, the individual expert remains indispensable who gives his personal opinion, and who bases his opinion to a considerable extent on intangible factors which would be difficult to formulate and which involve a large element of his own experience.

Let us now imagine ourselves in the place of the expert confronted with the problem of estimating the best value for some numerical parameter and consider the various factors which he must consider. I suppose the tactics of our critical expert will to a large extent be determined by what he considers to be the nature of the numerical parameter which concerns him. Perhaps the first question is whether the parameter is one capable of indefinitely great accuracy, or whether it is one for which the physical nature of the subject precludes the assignment of a single unique value and permits only the establishment of a range of values. There are two sorts of parameter which permit indefinitely great accuracy. There are in the first place conventional constants, such as the standard atmosphere, which is fixed absolutely, by fiat, mostly on the simple basis of general convenience. Then there are the mathematical constants such as pi or $\sqrt{2}$, which may be computed to any preassigned limit of precision and with regard to which there is com-

plete consensus of expert opinion. After the mathematical constants come the so-called universal physical constants, such as the velocity of light, or Planck's constant, or the charge on the electron. Such constants have to be determined by physical measurement, and since all physical measurements are subject to marginal experimental error, the "best" present values for these constants also are subject to necessary error. It might perhaps be thought that an estimate of the most probable present error could be obtained merely by a critical examination of the experimental material. Detailed examination will show, however, that this is not the case, but the present state of physical *theory* enters as well as the present status of instrumental technique. For at present physical theory recognizes two types of physical constant—those the determination of which is affected by the Heisenberg uncertainty relation so that mathematical absolute precision is never attainable, and measurements subject to no such limitation and ideally capable of indefinite and mathematical precision. An example of the first sort of constant would be a characteristic emission frequency, into the determination of which enter measurements of energy levels, and examples of the second sort are the velocity of light or the charge of the electron or Planck's h. Present theory treats these latter as exactly in the same class as the mathematical constants pi or $\sqrt{2}$. One need not question the present justifiability of dividing the fundamental constants into two classes in this way, to at least recognize that the present attitude toward these constants depends on present theoretical consensus, and that if at any time in the future wave mechanics should be modified, consensus about the nature of these constants would change also. To me it seems intuitively probable that eventually no physically determined constant can remain on the level of the mathematical constants, for some sort of indefiniteness in physical measurement will always remain, whether of the Heisenberg sort or not.

In general, hardly any sort of physical measurement is possible which does not involve some element of theory, and the tactics of the measurement may well be influenced by the present status of the theory. A striking example is afforded by the recognition of the possible existence of isotopes. I well remember the consternation of T. W. Richards, who had acquired a well-merited world-wide recognition for his precise determination of atomic weights, when confronted with the realization that all his so-careful determinations of what he thought were eternal verities were only more or less blurred-out averages with mostly a casual significance. Since then the recognition of the existence of isotopes has radically changed the entire technique of atomic weight measurement. Strictly, recognition of the existence of isotopes should alter the whole method of presenting the properties of actual materials. Ideally, no density of a metal, for example, should be tabulated without specifying the isotopic constitution of the metal. Going still further, a complete characterization of the density of a specific metal would demand tabulation of the density of all pure isotopes. This is probably at present an unnecessary and impossible ideal, but I think we have at least the right to demand in a table of densities some indication of the range to be expected because of possible variation in the isotopic constitution. The range to be expected will depend on theory to a certain extent. The same requirement applies to all the other physical parameters. Attention to this requirement will obviously materially add to the complexity of present tables.

The way in which the various physical parameters are tabulated is to a certain extent an indication of the present state of the theory. The indication may be negative, by omission. Thus if the most probable value of the velocity of light is listed without specifying the date at which it was determined, the presumption is that the velocity of light is regarded as an absolute constant, not subject to secular variation or to variation with the position of the solar system in the sidereal universe. Neglect of both these factors can doubtless be justified within present limits of error by our present theory of general relativity, but if we neglect to record these factors we are in effect gambling that our present general theory of relativity will continue to command universal assent (which can hardly be maintained of it even now). In general, a record of the date of any measurement would seem to be an integral part of the data pertaining to the measurement.

We turn now to consider more explicitly some of the factors which obscure the significance of the numerical values for some of the parameters of specific concrete substances, as distinguished from the so-called universal constants. We have already mentioned the uncertainty arising from ill-defined isotopic composition. Indefiniteness of isotopic composition is by no means the only source of uncertainty, but with increasing precision of measurement new and unsuspected sources of uncertainty are continually uncovered. The ideal here is to present measured parameters only for material that is well defined, and with scientific progress our specifications for "well-defined" become continually more exacting. It has always been recognized that certain materials as they naturally occur are not well defined, but nevertheless, because of their practical importance, some indication of the values to be expected should be tabulated. A typical example is that of the natural woods. It is, for example, of practical importance to list expected values for the densities of various woods, in spite of the fact that the densities of two specimens from different locations on the same tree are measurably different, and also variable with moisture content. Nevertheless, tables of the density of wood are of value. They become of greater value if the expected range of variation under conditions which are specified as exactly as possible is also tabulated. It is probably impossible to give any rigid formal specifications which we can justifiably demand the expert who prepares the table to follow. What the expert does will be colored by his personal experience, and it is to our advantage that it should be so. What we do have a right to demand is that the expert should take us into his confidence to the greatest possible extent, and explicitly formulate as many of the considerations which have guided him as he can.

The same sort of indefiniteness that prevails with regard to the density of wood is characteristic to a less pronounced degree of the densities of other materials. Twenty or more years ago it was the prevalent opinion that the only well-defined solid material was the perfect crystal, and an ideal density was the density of single crystal material. But we now recognize that even this does not take account of all the factors at present known. For in practice the perfect crystal does not exist, but all actual crystals have dislocations. The number of dislocations may be *relatively* small when the crystal is small, as in a "whisker," but in larger crystals the relative number becomes larger, and also becomes more nebulous, being a function of past history, and so is not a single valued temperature function. Not only is the measured density of a single crystal affected by the number of disloca-

tions, but it is also a function of absolute size and shape, when measurements are made accurately enough, because there are surface effects, and the mathematically precise lattice is distorted to a certain extent in the vicinity of the surface. For all these reasons the density of a single crystal of an isotopically pure substance is not well defined, and any critical table must give some indication of the range of variation to be expected.

If the material is "solid" but not a single crystal, there may be much greater uncertainties in its density. The density of a glass fluctuates through a range depending on its past thermal and mechanical history. If the possible previous mechanical stresses are increased by improvements in technique beyond limits previously thought attainable, fluctuations in density may be produced greater than at one time thought possible. Thus by subjecting quartz glass to a combination of shearing stress and hydrostatic pressure in excess of 50,000 kg/cm^2, a permanent increase of density of 17 per cent has been produced, the material remaining a perfect glass and without any trace of crystallinity as shown by X-ray analysis. With progress in technique other new factors are being found which are capable of producing permanent changes of density and therefore indefiniteness in the values which should be tabulated. There is, for example, the cross linkage in long chain polymers produced by electron bombardment, or the permanent changes in many metals produced after prolonged neutron bombardment in a pile. Or a photographic emulsion, made developable by a cosmic ray, has a density different from that of the unexposed emulsion. There is a logical puzzle here in dealing with this situation, for the fact that a cosmic ray has been received can usually be established only by developing the emulsion and thus irreversibly changing it.

When one burrows far enough into the microscopic structure of materials, one encounters an entirely new class of physical parameter—the parameter which can be measured but not controlled. Such parameters doubtless play a dominating role in irreversible processes. But science has hardly yet reached the stage where the parameters of substances undergoing irreversible changes can be adequately included in critical tables.

The densities of liquids are subject to some of the same uncertainties as the densities of solids, although to a less degree. In the first place, internal equilibrium between different molecular species requires time to achieve; sometimes the sluggishness may be so great as to make difficult a sharply defined value within present experimental limits. This is particularly the case for organic liquids which solidify to stable and unstable solid forms. A somewhat different source of uncertainty is shown by liquids in the neighborhood of the critical point with the vapor. It used to be thought, on the basis of van der Waals' equation of state, or of some similar equation, that the critical point was a single well defined point with definite values for the pressure and temperature. It is now realized that instead of a critical *point* we have a critical *region*, within which there are rapid fluctuations of small scale density, distributed about some sort of mean according to a probability function. The situation here adumbrates a situation that is eventually going to make trouble when the accuracy of physical measurement is pushed to the limit. It is universally recognized that in tabulating any such parameter as density the temperature must be specified. Now this "temperature" is the thermodynamic temperature, defined in terms of equilibrium.

But when we make our measurements fine enough we find that equilibrium is never attained, and that therefore this "temperature" strictly does not exist. Present theoretical consensus is inclined to the view that if we want to get down to bedrock fundamentals we must replace temperature by some sort of statistical average of the mechanical variables of kinetic theory. I personally am not sure that this sort of thing can be carried through "in principle," because it seems to me that there are unsolved difficulties in applying the mathematical concepts of probability theory to concrete situations. However, I think that everyone will agree that difficulties of this sort are mostly in the future, and that they will play an almost negligible role in constructing the practical tables which will ultimately result from our present deliberations.

Mention above of van der Waals' equation of state brings up another sort of consideration fundamental to the construction of critical tables. A complete reproduction of the density (or volume) of any substance demands that it be given as a function of all the variables which are known to affect it. In general, these variables include such things as electric or magnetic fields, but for most purposes we may consider only the most important variables, which may be taken to be simply pressure and temperature. Even with this reduction, the complete tabulation of density as a function of pressure and temperature would usually demand a prohibitive amount of space. In many cases the experimental material can be reproduced by one or the other of the many equations of state that have been proposed, and the purpose of the tables is fulfilled if the constants of the equations for such substances are tabulated instead of the full range of experimental material. But these equations of state are usually erected on some sort of theoretical foundation, and acceptance of a specific equation involves some sort of commitment with regard to what is to be expected when the experimental range is extended beyond that now attainable. In such cases, some discussion of the theoretical background of the equation is necessarily involved in any critical presentation of the material. It may be that our theoretical understanding is not sufficient to justify any *deduction* of an equation of state, but that nevertheless empirical equations can be found which reproduce the experimental results within experimental error. In some other cases, it may be that our theoretical understanding, although incomplete, may be good enough to suggest that a particular type of equation is better adapted than some other. In such cases, discussion of the reasons for preferring this or that type of equation should be included with the numerical material.

There is another sort of physical parameter which in practice is subject to much greater fluctuation than density, namely, breaking strength. Under the conditions of daily life, fracture is one of the most unmistakable things that can happen to one. Fracture is always catastrophic, irreversible, and there is no question when it has occurred. Nevertheless, in spite of the definiteness of the phenomenon, the parameters which characterize it are among the most indefinite of the parameters which it is the task of a critical table to tabulate. The breaking strength of different specimens of ostensibly the same material may vary in a way frustratingly discouraging when it comes to practical applications. Differences of constitution which for other phenomena may be unimportant may be decisive with regard to strength, such as slight differences of chemical composition, or of surface condition, or of past history. We are gradually acquiring a better understanding of the way in which

some of these obscure factors work, and also the ability to make specimens more reproducible with regard to fracture phenomena, but still our control is much less adequate than with respect to nearly all other physical parameters. Added to the practical difficulties there is a conceptual difficulty, for the concept of breaking strength becomes continuously less applicable when conditions are suitably varied, and eventually fails altogether. In particular, by raising the hydrostatic pressure to which a material is subjected the reduction of area before fracture may be increased indefinitely, so that eventually the material becomes incapable of tensile fracture and exhibits perfect plasticity. A further difficulty now appears, because in the plastic range the behavior of the material is by no means uniquely defined, but there are different kinds of plasticity, one or the other of which may be called "perfect" from one or another point of view. Furthermore, in the plastic range a formidable difficulty of principle appears, because no reversible displacements whatever are possible. This means that it is impossible, at least for the present, to even define an entropy for the material. On the other hand, it has to be assumed that the material has an entropy whenever it is characterized in thermodynamic terms, as it is for purposes of a table. In spite of all this, it is essential for practical purposes to attempt to give in a critical table some characterization of the behavior to be expected of ordinary materials in the range of fracture and plastic flow. It does not as yet appear even what are the best sorts of parameters in terms of which to attempt the characterization. We are here dependent to an unusual degree on the expert in charge of this part of the table.

This discussion has been concerned almost exclusively with parameters which may be characterized as "mechanical," namely, density and strength. But I think it is evident that somewhat similar considerations could have been advanced with regard to the other parameters which are the proper subject matter of critical tables, such as the electrical or optical parameters. Always, I think, the same sort of situation will be encountered, namely, that as the accuracy of measvrement is increased and as we become acquainted with a wider range of new phenomena, it will become increasingly difficult to exhaustively define the conditions that are necessary to obtain reproducible results. Never will it be possible to deal with a world like this in a completely formal manner, nor will it ever be possible to reduce the construction of a set of critical tables to a rigid set of rules. Always will inspired common sense remain the most important requirement of those who are to produce the tables.

* Introductory paper at the Gordon Conference on Information Processing for Critical Tables of Scientific Data, June 20, 1960, New Hampton, New Hampshire.

*THE SPECIAL CANONICAL GROUP**

By Julian Schwinger

HARVARD UNIVERSITY

Communicated August 30, 1960

This note is concerned with the further development and application of an operator group described in a previous paper.[1] It is associated with the quantum degree of freedom labeled $\nu = \infty$ which is characterized by the complementary pair of operators q, p with continuous spectra.[2] The properties of such a degree of freedom are obtained as the limit of one with a finite number of states, specifically given by a prime integer ν. We recall that unitary operators U and V obeying

$$U^\nu = V^\nu = 1, \; VU = e^{2\pi i/\nu} UV$$

define two orthonormal coordinate systems $\langle u^k|$ and $\langle v^k|$, where

$$u^k = v^k = e^{2\pi ik/\nu}$$

and

$$\langle u^k | v^l \rangle = \nu^{-1/2} e^{2\pi ikl/\nu}.$$

For any prime $\nu > 2$ we can choose the integers k and l to range from $-{}^1/_2(\nu - 1)$ to ${}^1/_2(\nu - 1)$, rather than from 0 to $\nu - 1$. An arbitrary state Ψ can be represented alternatively by the wave functions

$$\psi(u^k) = \langle u^k | \Psi, \; \psi(v^k) = \langle v^k | \Psi$$

where

$$\Psi^\dagger \Psi = \sum_k |\psi(u^k)|^2 = \sum_k |\psi(v^k)|^2,$$

and the two wave functions are reciprocally related by

$$\psi(u^k) = \sum_l \nu^{-1/2} \, e^{2\pi ikl/\nu} \, \psi(v^l)$$

$$\psi(v^l) = \sum_k \nu^{-1/2} \, e^{-2\pi ikl/\nu} \, \psi(u^k).$$

We now shift our attention to the Hermitian operators q, p defined by

$$U = e^{i\epsilon q}, \quad V = e^{i\epsilon p}, \quad \epsilon = (2\pi/\nu)^{1/2}$$

and the spectra

$$q', p' = \epsilon[-1/2(\nu - 1) \; .. \; 0 \; .. \; 1/2(\nu - 1)].$$

Furthermore, we redefine the wave functions so that

$$\psi(u^k) = \epsilon^{1/2} \psi(q'), \; \psi(v^k) = \epsilon^{1/2} \psi(p')$$

where $\epsilon = \Delta q' = \Delta p'$, the interval between adjacent eigenvalues. Then we have

$$\Psi^\dagger \Psi = \sum_{q'} \Delta q' |\psi(q')|^2 = \sum_{p'} \Delta p' |\psi(p')|^2$$

and

$$\psi(q') = \sum_{p'} \Delta p'(2\pi)^{-1/2} e^{iq'p'} \psi(p')$$

$$\psi(p') = \sum_{q'} \Delta q'(2\pi)^{-1/2} e^{-iq'p'} \psi(q').$$

As ν increases without limit the spectra of q and p become arbitrarily dense and the eigenvalues of largest magnitude increase indefinitely. Accordingly we must restrict all further considerations to that physical class of states, or physical subspace of vectors, for which the wave functions $\psi(q')$ and $\psi(p')$ are sufficiently well behaved with regard to continuity and the approach of the variable to infinity that a uniform transition to the limit $\nu = \infty$ can be performed, with the result

$$\Psi^\dagger\Psi = \int_{-\infty}^{\infty} dq' |\psi(q')|^2 = \int_{-\infty}^{\infty} dp' |\psi(p')|^2$$

and

$$\psi(q') = \int_{-\infty}^{\infty} dp' (2\pi)^{-1/2} e^{iq'p'} \psi(p')$$

$$\psi(p') = \int_{-\infty}^{\infty} dq' (2\pi)^{-1/2} e^{-iq'p'} \psi(q').$$

We shall not attempt here to delimit more precisely the physical class of states. Note however that the reciprocal relation between wave functions can be combined into

$$\psi(q') = \int_{-\infty}^{\infty} \frac{dp'}{2\pi} e^{iq'p'} \int_{-\infty}^{\infty} dq'' e^{-iq''p'} \psi(q'')$$

which must be an identity for wave functions of the physical class when the operations are performed as indicated. There will also be a class of functions $K(p', \epsilon)$, such that

$$K(p', \epsilon) \rightarrow \begin{cases} 1, & \epsilon \rightarrow 0 \\ 0, & |p'| \rightarrow \infty \end{cases}$$

and

$$\psi(q') = \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} \frac{dp'}{2\pi} e^{iq'p'} K(p', \epsilon) \int_{-\infty}^{\infty} dq'' e^{-iq''p'} \psi(q'')$$

$$= \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} dq'' \left[\int_{-\infty}^{\infty} \frac{dp'}{2\pi} K(p', \epsilon) e^{ip'(q'-q'')} \right] \psi(q'').$$

This is what is implied by the symbolic notation (Dirac)

$$\psi(q') = \int_{-\infty}^{\infty} dq'' \delta(q' - q'') \psi(q'')$$

$$\delta(q' - q'') = \int_{-\infty}^{\infty} \frac{dp'}{2\pi} e^{ip'(q'-q'')}$$

We shall also use the notation that is modeled on the discrete situation, with integrals replacing summations, as in

$$\langle q' | q'' \rangle = \int_{-\infty}^{\infty} \langle q' | p' \rangle dp' \langle p' | q'' \rangle = \delta(q' - q''),$$

with

$$\langle q' | p' \rangle = \langle p' | q' \rangle^* = (2\pi)^{-1/2} e^{iq'p'}.$$

There are other applications of the limit $\nu \to \infty$. The reciprocal property of the operators U and V is expressed by

$$\langle u^k | V = \langle u^{k+1} |, \ \langle v^k | U^{-1} = \langle v^{k+1} |$$

or

$$\langle q' | e^{i\epsilon p} = \langle q' + \epsilon |, \ \langle p' | e^{-i\epsilon q} = \langle p' + \epsilon |,$$

with an exception when q' or p' is the greatest eigenvalue, for then $q' + \epsilon$ or $p' + \epsilon$ is identified with the least eigenvalue, $-q'$ or $-p'$. We write these relations as wave function statements, of the form

$$\frac{1}{\epsilon}[\langle q' + \epsilon | \Psi - \langle q' | \Psi] = \langle q' | \frac{1}{\epsilon}(e^{i\epsilon p} - 1)\Psi$$

and

$$\frac{1}{\epsilon}[\langle p' + \epsilon | \Psi - \langle p' | \Psi] = \langle p' | \frac{1}{\epsilon}(e^{-i\epsilon q} - 1)\Psi.$$

In the transition to the limit $\nu = \infty$, the subspace of physical vectors Ψ is distinguished by such properties of continuity and behavior at infinity of the wave functions that the left-hand limits $\epsilon \to 0$, exist as derivatives of the corresponding wave function. We conclude, for the physical class of states, that

$$\frac{1}{i}\frac{\partial}{\partial q'}\langle q' | \Psi = \langle q' | p\Psi$$

and

$$i\frac{\partial}{\partial p'}\langle p' | \Psi = \langle p' | q\Psi.$$

It will also be evident from this application of the limiting process to the unitary operators $\exp(i\epsilon q)$, $\exp(i\epsilon p)$ that a restriction to a physical subspace is needed for the validity of the commutation relation

$$[q, p] = i.$$

The elements of an orthonormal operator basis are given by

$$\nu^{-1/2} e^{\pi i m n/\nu} U^m V^n = \nu^{-1/2} e^{-(\pi i m n/\nu)} V^n U^m$$

or $\nu^{-1/2} U(q'p')$, with

$$U(q'p') = e^{1/2 i p'q'} e^{ipq'} e^{-ip'q} = e^{-1/2 i p'q'} e^{-ip'q} e^{ipq'}$$

$$= (\nu \to \infty)\, e^{i(pq' - p'q)}.$$

Thus, since $\nu^{-1} = \Delta q' \Delta p'/2\pi$, we have, for an arbitrary function $f(q'p')$ of the discrete variables q', p',

$$\sum_{q''p''} \frac{\Delta q'' \Delta p''}{2\pi} tr\, U(q'p')^\dagger U(q''p'') f(q''p'') = f(q'p').$$

In the limit $\nu = \infty$ there is a class of functions $f(q'p')$ such that

$$tr\ U(q'p')^\dagger \int_{-\infty}^{\infty} \frac{dq''dp''}{2\pi}\ U(q''p'')f(q''p'') = f(q'p'),$$

which we express symbolically by

$$tr\ U(q'p')^\dagger U(q''p'') = 2\pi\delta(q' - q'')\delta(p' - p'').$$

In particular,

$$tr\ e^{i(pq'-p'q)} = 2\pi\delta(q')\delta(p')$$
$$= \int_{-\infty}^{\infty} \frac{dqdp}{2\pi}\ e^{i(pq'-p'q)}$$

where q and p on the right-hand side are numerical integration variables.

The completeness of the operator basis $U(q'p')$ is expressed by

$$\int_{-\infty}^{\infty} \frac{dq'dp'}{2\pi}\ U(q'p')\ X\ U(q'p')^\dagger = 1\ tr\ X.$$

The properties of the $U(q'p')$ basis are also described, with respect to an arbitrary discrete operator basis $X(\alpha)$, by

$$\int_{-\infty}^{\infty} \frac{dq'dp'}{2\pi}\ \langle\alpha|\,q'p'\rangle\langle q'p'|\,\alpha'\rangle = \delta(\alpha, \alpha')$$

and

$$\sum_{\alpha} \langle q'p'|\,\alpha\rangle\langle\alpha|\,q''p''\rangle = 2\pi\delta(q' - q'')\delta(p' - p'').$$

When X is given as $F(q, p)$, the operations of the special canonical group can be utilized to bring the completeness expression into the form

$$\int_{-\infty}^{\infty} \frac{dq'dp'}{2\pi}\ F(q + q', p + p') = 1\ tr\ F(q, p).$$

This operator relation implies a numerical one if it is possible to order $F(q, p)$, so that all q operators stand to the left, for example, of the operator p: $F(q; p)$. Then the evaluation of the $\langle q' = 0|\quad|p' = 0\rangle$ matrix element gives

$$tr\ F(q, p) = \int_{-\infty}^{\infty} \frac{dqdp}{2\pi}\ F(q; p)$$

which result also applies to a system with n continuous degrees of freedom if it is understood that

$$\frac{dqdp}{2\pi} = \prod_{k=1}^{n} \frac{dq_k dp_k}{2\pi}.$$

As an example of this ordering process other than the one already given by $tr\ U(q'p')$, we remark that

$$e^{-1/2(q^2+p^2)\beta} = (\cosh\beta)^{-\frac{1}{2}}\ e^{-\frac{1}{2}q^2\tanh\beta}\ e^{iq;\ p(\mathrm{sech}\ \beta-1)}\ e^{-1/2p^2\tanh\beta}$$

where

$$e^{a;\ b} = \sum_{0}^{\infty} \frac{1}{n!}\ a^n b^n,$$

and therefore

$$tr\ e^{-1/2(q^2+p^2)\beta} = \int_{-\infty}^{\infty} \frac{dqdp}{2\pi} (\cosh \beta)^{-1/2}\ e^{-1/2(q^2+p^2)\tanh \beta}\ e^{iqp\ (\mathrm{sech}\ \beta - 1)}$$

$$= \frac{1}{2 \sinh {}^1/_2\beta} = \sum_{n=0}^{\infty} e^{-(n+1/2)\beta},$$

which reproduces the well-known non-degenerate spectrum of the operator ${}^1/_2(q^2 + p^2)$.

Now we shall consider the construction of finite special canonical transformations from a succession of infinitesimal ones, as represented by the variation of a parameter τ. Let the generator of the transformation associated with $\tau \to \tau + d\tau$ be

$$d\tau G_s = d\tau(qP - pQ)$$

where $Q(\tau)$ and $P(\tau)$ are arbitrary numerical functions of τ. That is, the infinitesimal transformation is

$$q(\tau + d\tau) - q(\tau) = -\frac{1}{i}[q, d\tau G_s] = d\tau Q(\tau)$$

$$p(\tau + d\tau) - p(\tau) = -\frac{1}{i}[p, d\tau G_s] = d\tau P(\tau),$$

which implies the finite transformation

$$q(\tau_1) - q(\tau_2) = \int_{\tau_2}^{\tau_1} d\tau Q(\tau)$$

$$p(\tau_1) - p(\tau_2) = \int_{\tau_2}^{\tau_1} d\tau P(\tau).$$

Some associated transformation functions are easily constructed. We have

$$\langle \tau + d\tau| = \langle \tau|\,[1 + id\tau G(x(\tau), \tau)]$$

or

$$\frac{1}{i}\frac{\partial}{\partial \tau}\langle \tau| = \langle \tau|\,[q(\tau)P(\tau) - p(\tau)Q(\tau)],$$

and therefore

$$\frac{1}{i}\frac{\partial}{\partial \tau}\langle q'\tau|\,p'\tau_2\rangle$$

$$= \langle q'\tau|\,[q'P(\tau) - (p' + \int_{\tau_2}^{\tau} d\tau' P(\tau'))Q(\tau)]\,|p'\tau_2\rangle$$

which, in conjunction with the initial condition

$$\tau = \tau_2\colon\ \langle q'\tau|\,p'\tau_2\rangle = \langle q'|\,p'\rangle = (2\pi)^{-1/2} e^{iq'p'}$$

gives

$$\langle q'\tau_1|\,p'\tau_2\rangle^{QP} = (2\pi)^{-1/2} \exp\,[i(q'p' + q' \int_{\tau_2}^{\tau_1} d\tau P(\tau) - p' \int_{\tau_2}^{\tau_1} d\tau Q(\tau) - \int_{\tau_2}^{\tau_1} d\tau d\tau' Q(\tau)\eta_+(\tau - \tau')P(\tau'))]$$

with

$$\eta_+(\tau - \tau') = \begin{cases} 1, & \tau > \tau' \\ 0, & \tau < \tau' \end{cases}.$$

From this result we derive

$$\langle q'\tau_1|\, q''\tau_2\rangle^{QP} = \int_{-\infty}^{\infty} \langle q'\tau_1|\, p'\tau_2\rangle^{QP} dp' \langle p'|\, q''\rangle = \delta(q' - q'' - \int_{\tau_2}^{\tau_1} d\tau Q(\tau))\; e^{iq'\int d\tau P}\, e^{-i\int d\tau d\tau' Q\eta_+ P},$$

and

$$\langle p'\tau_1|\, p''\tau_2\rangle^{QP} = \int_{-\infty}^{\infty} \langle p'|\, q'\rangle dq' \langle q'\tau_1|\, p''\tau_2\rangle^{QP} = \delta(p' - p'' - \int_{\tau_2}^{\tau_1} d\tau P(\tau))\; e^{-ip''\int d\tau Q}\, e^{-i\int d\tau d\tau' Q\eta_+ P}$$

or, alternatively,

$$\delta(p' - p'' - \int d\tau P)\; e^{-ip'\int d\tau Q}\, e^{i\int d\tau d\tau' P\eta_+ Q},$$

on using the fact that

$$\eta_-(\tau - \tau') = 1 - \eta_+(\tau - \tau') = \eta_+(\tau' - \tau).$$

These transformation functions can also be viewed as matrix elements of the unitary operator, an element of the special canonical group, that produces the complete transformation. That operator, incidentally, is

$$\exp[i(q\int_{\tau_2}^{\tau_1} d\tau P(\tau) - p\int_{\tau_2}^{\tau_1} d\tau Q(\tau) - {}^1/_2 \int_{\tau_2}^{\tau_1} d\tau d\tau' Q(\tau)\epsilon(\tau - \tau')P(\tau'))]$$

where $\epsilon(\tau - \tau')$ is the odd step function

$$\epsilon = \eta_+ - \eta_-.$$

We can compute the trace of a transformation function, regarded as a matrix, and this will equal the trace of the associated unitary operator provided the otherwise arbitrary representation is not an explicit function of τ. Thus

$$tr\langle\tau_1|\, \tau_2\rangle^{QP} = \int_{-\infty}^{\infty} dq' \langle q'\tau_1|\, q'\tau_2\rangle^{QP} = \int_{-\infty}^{\infty} dp' \langle p'\tau_1|\, p'\tau_2\rangle^{QP} = 2\pi\delta\,(\int_{\tau_2}^{\tau_1} d\tau Q(\tau))\delta(\int_{\tau_2}^{\tau_1} d\tau P(\tau))\; e^{-i\int d\tau d\tau' Q\eta_+ P}$$

where, in view of the delta function factors, $\eta_+(\tau - \tau')$ can be replaced by other equivalent functions, such as $\eta_+ - {}^1/_2 = {}^1/_2\epsilon$, or ${}^1/_2\epsilon(\tau - \tau') - (\tau - \tau')/T$, with $T = \tau_1 - \tau_2$. The latter choice has the property of giving a zero value to the double integral whenever $Q(\tau)$ or $P(\tau)$ is a constant. As an operator statement, the trace formula is the known result

$$tr\; e^{i(pq' - p'q)} = 2\pi\delta(q')\delta(p').$$

It is important to recognize that the trace, which is much more symmetrical than any individual transformation function, also implies specific transformation functions. Thus, let us make the substitutions

$$\begin{aligned} Q(\tau) &\to Q(\tau) - (q' - q'')\delta(\tau - \tau_1 + \epsilon T) \\ P(\tau) &\to P(\tau) + p'\delta(\tau - \tau_2 - \epsilon T) \end{aligned} \quad , \epsilon \to +0$$

and then indicate the effect of the additional localized transformations by the equivalent unitary operators, which gives

$$tr\langle\tau_1|\; e^{i(q' - q'')p(\tau_1)}\, e^{ip'q(\tau_2)}\,|\tau_2\rangle^{QP} = \int_{-\infty}^{\infty} dq \langle q + q' - q'' \qquad \tau_1|\, q\tau_2\rangle^{QP}\, e^{ip'q}.$$

Accordingly, if we also multiply by $\exp(-ip'q'')$ and integrate with respect to $dp'/2\pi$, what emerges is $\langle q'\tau_1|\, q''\tau_2\rangle^{QP}$, as we can verify directly.

We shall find it useful to give an altogether different derivation of the trace formula. First note that

$$tr\langle\tau_1|\,(q(\tau_1) - q(\tau_2))\,|\tau_2\rangle = \int_{-\infty}^{\infty} dq'\langle q'\tau_1|\,(q' - q')\,|q'\tau_2\rangle = 0$$

and similarly

$$tr\langle\tau_1|\,(p(\tau_1) - p(\tau_2))\,|\tau_2\rangle = 0$$

which is a property of periodicity over the interval $T = \tau_1 - \tau_2$. Let us, therefore, represent the operators $q(\tau)$, $p(\tau)$ by the Fourier series

$$q(\tau) = q_0 + \sum_1^{\infty} (\pi n)^{-1/2}\left[q_n \cos\frac{2\pi n}{T}(\tau - \tau_2) + q_{-n}\sin\frac{2\pi n}{T}(\tau - \tau_2)\right]$$

$$p(\tau) = p_0 + \sum_1^{\infty} (\pi n)^{-1/2}\left[-p_n \sin\frac{2\pi n}{T}(\tau - \tau_2) + p_{-n}\cos\frac{2\pi n}{T}(\tau - \tau_2)\right],$$

where the coefficients are so chosen that the action operator for an arbitrary special canonical transformation

$$W_{12} = \int_{\tau_2}^{\tau_1} [{}^1/_2(pdq - dpq) + d\tau(qP - pQ)]$$

acquires the form

$$W_{12} = \sum_{-\infty}^{\infty}{}' p_n q_n + \sum_{-\infty}^{\infty} (q_n P_n - p_n Q_n).$$

Here the dash indicates that the term $n = 0$ is omitted, and

$$Q_0 = \int_{\tau_2}^{\tau_1} d\tau Q(\tau),\quad P_0 = \int_{\tau_2}^{\tau_1} d\tau P(\tau)$$

$$Q_{n,-n} = (\pi n)^{-1/2}\int_{\tau_2}^{\tau_1} d\tau Q(\tau)(-\sin,\cos)\left(\frac{2\pi n}{T}(\tau - \tau_2)\right)$$

$$P_{n,-n} = (\pi n)^{-1/2}\int_{\tau_2}^{\tau_1} d\tau P(\tau)(\cos,\sin)\left(\frac{2\pi n}{T}(\tau - \tau_2)\right).$$

The action principle for the trace is

$$\delta\, tr\,\langle\tau_1|\tau_2\rangle = i\, tr\,\langle\tau_1|\,\delta[W_{12}]\,|\tau_2\rangle$$

and the principle of stationary action asserts that

$$Q_0\, tr\,\langle\tau_1|\tau_2\rangle^{QP} = P_0\, tr\,\langle\tau_1|\tau_2\rangle^{QP} = 0,$$

together with

$$tr\,\langle\tau_1|\,(q_n - Q_n)\,|\tau_2\rangle^{QP} = tr\,\langle\tau_1|\,(p_n + P_n)\,|\tau_2\rangle^{QP} = 0.$$

The first of these results implies that the trace contains the factors $\delta(Q_0)$ and $\delta(P_0)$. The dependence upon Q_n, P_n, $n \neq 0$, is then given by the action principle as

$$\frac{\partial}{\partial Q_n}(tr) = i\, tr\,\langle\tau_1|\,(-p_n)\,|\tau_2\rangle = i\,P_n\,(tr)$$

$$\frac{\partial}{\partial P_n}(tr) = i\, tr\,\langle\tau_1|\,q_n\,|\tau_2\rangle = iQ_n(tr).$$

Therefore

$$tr\,\langle\tau_1|\,\tau_2\rangle^{QP} = 2\pi\delta(Q_0)\delta(P_0)\,e^{i\sum_{-\infty}^{\infty}{}' Q_nP_n}$$

where the factor of 2π is supplied by reference to the elementary situation with constant $Q(\tau)$ and $P(\tau)$. We note, for comparison with previous results, that

$$Q_nP_n + Q_{-n}P_{-n} = -\frac{1}{\pi n}\int_{\tau_2}^{\tau_1} d\tau d\tau' Q(\tau)\sin\left(\frac{2\pi n}{T}(\tau-\tau')\right)P(\tau')$$

and

$$\sum_1^\infty \frac{1}{\pi n}\sin\frac{2\pi n}{T}(\tau-\tau') = {}^1\!/_2\,\epsilon(\tau-\tau') - \frac{\tau-\tau'}{T}.$$

The new formula for the trace can be given a uniform integral expression by using the representation

$$e^{iQP} = \int_{-\infty}^{\infty}\frac{dqdp}{2\pi}\,e^{i(pq+qP-pQ)}$$

for now we can write

$$tr\,\langle\tau_1|\,\tau_2\rangle^{QP} = \int d[q,\,p]\,e^{iW[q,\,p]},$$

where

$$d[q,\,p] = \prod_{-\infty}^{\infty}\frac{dq_ndp_n}{2\pi}$$

and $W[q, p]$ is the numerical function formed in the same way as the action operator,

$$W[q,\,p] = \sum_{-\infty}^{\infty}{}'\,p_nq_n + \sum_{-\infty}^{\infty}(q_nP_n - p_nQ_n).$$

Alternatively, we can use the Fourier series to define the numerical functions $q(\tau)$, $p(\tau)$. Then

$$W[q,\,p] = \int_{\tau_2}^{\tau_1}[{}^1\!/_2(pdq - dpq) + d\tau(qP - pQ)]$$

and $d[q, p]$ appears as a measure in the quantum phase space of the functions $q(\tau)$, $p(\tau)$.

It is the great advantage of the special canonical group that these considerations can be fully utilized in discussing arbitrary additional unitary transformations, as described by the action operator

$$W_{12} = \int_{\tau_2}^{\tau_1}[{}^1\!/_2(pdq - dpq) + d\tau(qP - pQ) + d\tau G(x(\tau),\,\tau)].$$

First, let us observe how an associated transformation function $\langle\tau_1|\,\tau_2\rangle_a^{QP}$ depends upon the arbitrary functions $Q(\tau)$, $P(\tau)$. The action principle asserts that

$$\delta QP\,\langle\tau_1|\,\tau_2\rangle = i\langle\tau_1|\int_{\tau_2}^{\tau_1}d\tau(q\delta P - p\delta Q)|\,\tau_2\rangle = i\int_{\tau_2}^{\tau_1}d\tau[\delta P(\tau)\langle\tau_1|\,q(\tau)|\,\tau_2\rangle - \delta Q(\tau)\langle\tau_1|\,p(\tau)|\,\tau_2\rangle]$$

which we express by the notation

$$-i\frac{\delta}{\delta P(\tau)}\langle\tau_1|\,\tau_2\rangle = \langle\tau_1|\,q(\tau)|\,\tau_2\rangle$$

$$i\frac{\delta}{\delta Q(\tau)}\langle\tau_1|\,\tau_2\rangle = \langle\tau_1|\,p(\tau)|\,\tau_2\rangle.$$

More generally, if $F(\tau')$ is an operator function of $x(\tau')$ but not of Q, P we have

$$\langle\tau_1|\,F(\tau')|\,\tau_2\rangle_G{}^{QP} = \langle\tau_1|\,\tau'\rangle_G{}^{QP} \times \langle\tau'|\,F(\tau')|\,\tau'\rangle \times \langle\tau'|\,\tau_2\rangle_G{}^{QP},$$

where $|\,\tau'\rangle \times \langle\tau'|$ symbolizes the summation over a complete set of states, and therefore

$$-i\frac{\delta}{\delta P(\tau)}\langle\tau_1|\,F(\tau')|\,\tau_2\rangle = \langle\tau_1|\,(q(\tau)F(\tau'))_0|\,\tau_2\rangle$$

$$i\frac{\delta}{\delta Q(\tau)}\langle\tau_1|\,F(\tau')|\,\tau_2\rangle = \langle\tau_1|\,(p(\tau)F(\tau'))_0|\,\tau_2\rangle.$$

Here $(\quad)_0$ is an ordered product that corresponds to the sense of progression from τ_2 to τ_1. If τ follows τ' in sequence, the operator function of τ stands to the left, while if τ precedes τ' the associated operator appears on the right. This description covers the two algebraic situations: $\tau_1 > \tau_2$, where we call the ordering positive, $(\quad)_+$, and $\tau_1 < \tau_2$, which produces negative ordering, $(\quad)_-$. For the moment take $\tau_1 > \tau_2$ and compare

$$i\frac{\delta}{\delta Q(\tau)}\langle\tau_1|\,q(\tau+0)|\,\tau_2\rangle = \langle\tau_1|\,q(\tau)p(\tau)|\,\tau_2\rangle$$

with

$$i\frac{\delta}{\delta Q(\tau)}\langle\tau_1|\,q(\tau-0)|\,\tau_2\rangle = \langle\tau_1|\,p(\tau)q(\tau)|\,\tau_2\rangle.$$

The difference of these expressions,

$$i\frac{\delta}{\delta Q(\tau)}\langle\tau_1|\,[q(\tau+0)-q(\tau-0)]|\,\tau_2\rangle = \langle\tau_1|\,[q(\tau),\,p(\tau)]|\,\tau_2\rangle,$$

refers on one side to the noncommutativity of the complementary variables q and p, and on the other to the "equation of motion" of the operator $q(\tau)$. According to the action principle

$$\frac{dq}{d\tau}+\frac{\partial G}{\partial p} = Q$$

and therefore

$$q(\tau+0)-q(\tau-0) = \operatorname*{Lim}_{\epsilon\to 0}\int_{\tau-\epsilon}^{\tau+\epsilon}d\tau'\left[-\frac{\partial G}{\partial p}+Q(\tau')\right]$$

which yields the expected result,

$$i\frac{\delta}{\delta Q(\tau)}\langle\tau_1|\,[q(\tau+0)-q(\tau-0)]|\,\tau_2\rangle = i\langle\tau_1|\,\tau_2\rangle.$$

Thus, through the application of the special canonical group, we obtain functional differential operator representations for all the dynamical variables. The general statement is

$$\langle \tau_1 | F(q, p)_o | \tau_2 \rangle = F\left(-i\frac{\delta}{\delta P}, i\frac{\delta}{\delta Q}\right)\langle \tau_1 | \tau_2 \rangle$$

where $F(q, p)_o$ is an ordered function of the $q(\tau)$, $p(\tau)$ throughout the interval between τ_2 and τ_1, and, as the simple example of $q(\tau)p(\tau)$ and $p(\tau)q(\tau)$ indicates, the particular order of multiplication for operators with a common value of τ must be reproduced by a suitable limiting process from different τ values.

The connection with the previous considerations emerges on supplying G with a variable factor λ. For states at τ_1 and τ_2 that do not depend explicitly upon λ we use the action principle to evaluate

$$\frac{\partial}{\partial\lambda}\langle \tau_1 | \tau_2 \rangle_{\lambda G}{}^{QP} = i\langle \tau_1 | \int_{\tau_2}^{\tau_1} d\tau G(qp\tau) | \tau_2 \rangle_{\lambda G}{}^{QP} = i\int_{\tau_2}^{\tau_1} d\tau G\left(-i\frac{\delta}{\delta P}, i\frac{\delta}{\delta Q}, \tau\right)\langle \tau_1 | \tau_2 \rangle_{\lambda G}{}^{QP}.$$

The formal expression that gives the result of integrating this differential equation from $\lambda = 0$ to $\lambda = 1$ is

$$\langle \tau_1 | \tau_2 \rangle_G{}^{QP} = \exp\left[i\int_{\tau_2}^{\tau_1} d\tau G\left(-i\frac{\delta}{\delta P}, i\frac{\delta}{\delta Q}, \tau\right)\right]\langle \tau_1 | \tau_2 \rangle^{QP}$$

where the latter transformation function is that for $\lambda = 0$ and therefore refers only to the special canonical group. An intermediate formula, corresponding to $G = G_1 + G_2$, contains the functional differential operator constructed from G_1 acting on the transformation function associated with G_2. The same structure applies to the traces of the transformation functions. If we use the integral representation for the trace of the special canonical transformation function, and perform the differentiations under the integration sign, we obtain the general integral formula[3]

$$tr\,\langle \tau_1 | \tau_2 \rangle_G{}^{QP} = \int d[q, p] e^{iW[q, p]}$$

Here the action functional $W[q, p]$ is

$$W[q, p] = \int_{\tau_2}^{\tau_1} [{}^1/_2(pdq - dpq) + d\tau(q(\tau)P(\tau) - p(\tau)Q(\tau)) + d\tau G(q(\tau)p(\tau)\tau)],$$

which is formed in essentially the same way as the action operator W_{12}, the multiplication order of noncommutative operator factors in G being replaced by suitable infinitesimal displacements of the parameter τ. The Hermitian operator G can always be constructed from symmetrized products of Hermitian functions of q and of p, and the corresponding numerical function is real. Thus the operator ${}^1/_2\{f_1(q(\tau), f_2(p(\tau)\}$ is represented by ${}^1/_2(f_1(q(\tau + \epsilon)) + f_1(q(\tau - \epsilon)))f_2(p(\tau))$, for example. One will expect to find that this averaged limit, $\epsilon \to 0$, implies no more than the direct use of $f_1(q(\tau)f_2(p(\tau))$, although the same statement is certainly not true of either term containing ϵ. Incidentally, it is quite sufficient to construct

the action from pdq, for example, rather than the more symmetrical version, in virtue of the periodicity.

As a specialization of this trace formula, we place $Q = P = 0$ and consider the class of operators G that do not depend explicitly upon τ. Now we are computing

$$tr\, e^{iTG} = \int d[q, p] e^{iW[q, p]}$$

$$W[q, p] = \int_0^T (pdq + d\tau G),$$

in which we have used the possibility of setting $\tau_2 = 0$. A simple example is provided by $G = {}^1/_2(p^2 + q^2)$, where

$$W = {}^1/_2T(p_0^2 + q_0^2) + \sum_{-\infty}^{\infty}{}' \left[p_n q_n + \frac{T}{4\pi|n|} (p_n^2 + q_n^2) \right]$$

and

$$tr\, e^{iT1/2(p^2+q^2)} = \int_{-\infty}^{\infty} \frac{dqdp}{2\pi} e^{iT1/2(p^2+q^2)} \times \prod_1^{\infty} \left[\int_{-\infty}^{\infty} \frac{dqdp}{2\pi} e^{1/2i(2qp+T/2\pi n(p^2+q^2))} \right]^2 =$$

$$\frac{i}{T} \prod_1^{\infty} \frac{1}{1 - \left(\frac{T}{2\pi n}\right)^2} = \frac{i}{2 \sin {}^1/_2T}.$$

This example also illustrates the class of Hermitian operators with spectra that are bounded below and for which the trace of exp (iTG) continues to exist on giving T a positive imaginary component, including the substitution $T \to i\beta$, $\beta > 0$. The trace formula can be restated for the latter situation on remarking that the Fourier series depend only upon the variable $(\tau - \tau_2)/T = \lambda$, which varies from 0 to 1, and therefore

$$tr\, e^{-\beta G} = \int d[q, p] e^{-w[q, p]}$$

$$w[q, p] = \int_0^1 d\lambda \left[-ip \frac{dq}{d\lambda} + \beta G(q(\lambda)p(\lambda)) \right].$$

Another property of the example is that the contributions to the trace of all Fourier coefficients except $n = 0$ tend to unity for sufficiently small T or β. This is also true for a class of operators of the form $G = {}^1/_2p^2 + f(q)$. By a suitable translation of the Fourier coefficients for $p(\lambda)$ we can write $w[q, p]$ as

$$w = \int_0^1 d\lambda \left[{}^1/_2\beta p(\lambda)^2 + \frac{1}{2\beta} \left(\frac{dq}{d\lambda} \right)^2 + \beta f(q(\lambda)) \right].$$

For sufficiently small β, the term involving $dq/d\lambda$, to which all the Fourier coefficients of $q(\lambda)$ contribute except q_0, will effectively suppress these Fourier coefficients provided appropriate restrictions are imposed concerning singular points in the neighborhood of which $f(q)$ acquires arge negative values. Then $f(q(\lambda) \sim f(q_0))$ and we can reduce the integrations to just the contribution of q_0 and p_0, as expressed by

$$tr\, e^{-\beta G(q, p)} \sim \int_{-\infty}^{\infty} \frac{dqdp}{2\pi} e^{-\beta G(q, p)}.$$

Comparison with the previously obtained trace formula involving the ordering of operators shows that the noncommutativity of q and p is not significant in this limit. Thus we have entered the classical domain, where the incompatibility of physical properties at the microscopic level is no longer detectible. Incidentally, a first correction to the classical trace evaluation, stated explicitly for one degree of freedom, is

$$tr\, e^{-\beta[1/2p^2+f(q)]} \sim \int_{-\infty}^{\infty} \frac{dq}{(2\pi\beta)^{1/2}}\, e^{-\beta f(q)}\, \frac{{}^1/_2\beta(f''(q))^{1/2}}{\sinh {}^1/_2\beta(f''(q))^{1/2}},$$

which gives the exact value when $f(q)$ is a positive multiple of q^2.

Another treatment of the general problem can be given on remarking that the equations of motion implied by the stationary action principle,

$$\frac{dq}{d\tau} + \frac{\partial G}{\partial p} = Q, \quad \frac{dp}{d\tau} - \frac{\partial G}{\partial q} = P,$$

can be represented by functional differential equations[4]

$$\left[\frac{d}{d\tau}(-i)\frac{\delta}{\delta P(\tau)} + \frac{\partial G}{\partial p}\left(-i\frac{\delta}{\delta P}, i\frac{\delta}{\delta Q}\right) - Q(\tau)\right]\langle\tau_1|\,\tau_2\rangle_G^{QP} = 0$$

$$\left[\frac{d}{d\tau} i\frac{\delta}{\delta Q(\tau)} - \frac{\partial G}{\partial q}\left(-i\frac{\delta}{\delta P}, i\frac{\delta}{\delta Q}\right) - P(\tau)\right]\langle\tau_1|\,\tau_2\rangle_G^{QP} = 0.$$

These equations are valid for any such transformation function. The trace is specifically distinguished by the property of periodicity,

$$\left(\frac{\delta}{\delta Q(\tau_1)} - \frac{\delta}{\delta Q(\tau_2)}\right) tr\,\langle\tau_1|\,\tau_2\rangle_G^{QP} = \left(\frac{\delta}{\delta P(\tau_1)} - \frac{\delta}{\delta P(\tau_2)}\right)(tr) = 0$$

which asserts that the trace depends upon $Q(\tau)$, $P(\tau)$ only through the Fourier coefficients Q_n, P_n, and that the functional derivatives can be interpreted by means of ordinary derivatives:

$$\frac{\delta}{\delta Q(\tau)} = \frac{\partial}{\partial Q_0} + \sum_1^\infty (\pi n)^{-1/2}\left[-\sin\frac{2\pi n}{T}(\tau-\tau_2)\frac{\partial}{\partial Q_n} + \cos\frac{2\pi n}{T}(\tau-\tau_2)\frac{\partial}{\partial Q_{-n}}\right]$$

$$\frac{\delta}{\delta P(\tau)} = \frac{\partial}{\partial P_0} + \sum_1^\infty (\pi n)^{-1/2}\left[\cos\frac{2\pi n}{T}(\tau-\tau_2)\frac{\partial}{\partial P_n} + \sin\frac{2\pi n}{T}(\tau-\tau_2)\frac{\partial}{\partial P_{-n}}\right].$$

Now introduce the functional differential operator that is derived from the numerical action function $W_G[q, p]$ which refers only to the transformation generated by G, namely

$$W_G\left[-i\frac{\delta}{\delta P}, i\frac{\delta}{\delta Q}\right] = \int_{\tau_2}^{\tau_1}\left[\frac{\delta}{\delta Q}\,d\,\frac{\delta}{\delta P} + d\tau G\left(-i\frac{\delta}{\delta P}, i\frac{\delta}{\delta Q}, \tau\right)\right],$$

and observe that the differential equations are given by

$$(i[W_G, Q(\tau)] + Q(\tau))\, tr\,\langle\tau_1|\,\tau_2\rangle_G^{QP} = 0$$

$$(i[W_G, P(\tau)] + P(\tau))\, tr\,\langle\tau_1|\,\tau_2\rangle_G^{QP} = 0,$$

or by

$$e^{iW_G}Q(\tau)e^{-iW_G}(tr) = e^{iW_G}P(\tau)e^{-iW_G}(tr) = 0.$$

The latter form follows from the general expansion

$$e^A Be^{-A} = B + [A, B] + {}^1/_2![A, [A, B]] + \ldots$$

on noting that $[W_G, Q(\tau)]$, for example, is constructed entirely from differential operators and is commutative with the differential operator W_G. Accordingly,

$$Q(\tau)e^{-iW_G}(tr) = P(\tau)e^{-iW_G}(tr) = 0,$$

which asserts that $\exp(-iW_G)(tr)$ vanishes when multiplied by any of the Fourier coefficients Q_n, P_n and therefore contains a delta function factor for each of these variables. We conclude that

$$tr\,\langle\tau_1|\,\tau_2\rangle_G{}^{QP} = e^{iW_G[-i\,\delta/\delta P,\; i\,\delta/\delta Q]}\,\delta[Q, P]$$

where, anticipating the proper normalization constants,

$$\delta[Q, P] = \prod_{-\infty}^{\infty} 2\pi\delta(Q_n)\delta(P_n).$$

A verification of these factors can be given by placing $G = 0$, which returns us to the consideration of the special canonical transformations. In this procedure we encounter the typical term

$$e^{i\,\partial/\partial Q\;\partial/\partial P}\,2\pi\delta(Q)\delta(P) = e^{iQP}$$

the proof of which follows from the remarks that

$$\left(Q + i\,\frac{\partial}{\partial P}\right)e^{i\,\partial/\partial Q\;\partial/\partial P} = e^{i\,\partial/\partial Q\;\partial/\partial P}\,Q$$

and

$$\int_{-\infty}^{\infty}\frac{dP}{2\pi}\,e^{i\,\partial/\partial Q\;\partial/\partial P}\,2\pi\delta(Q)\delta(P) = \delta(Q).$$

The result is just the known form of the transformation function trace

$$tr\,\langle\tau_1|\,\tau_2\rangle^{QP} = 2\pi\delta(Q_0)\delta(P_0)e^{\,i-\sum_{0}^{\infty}{}' Q_nP_n}$$

When integral representations are inserted for each of the delta function factors in $\delta[Q, P]$, we obtain

$$\delta[Q, P] = \prod_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{dq_n dp_n}{2\pi}\,e^{i(q_nP_n - p_nQ_n)} = \int d[q, p]\,e^{i\int_{\tau_2}^{\tau_1} d\tau(q(\tau)P(\tau) - p(\tau)Q(\tau))}$$

and the consequence of performing the differentiations in W_G under the integration signs is[5]

$$tr\,\langle\tau_1|\,\tau_2\rangle_G{}^{QP} = \int d[q, p]\,e^{\,iW[q,\,p]}$$

where the action function $W[q, p]$ now includes the special canonical transformation described by Q and P.

We shall also write this general integral formula as

$$tr\,\langle\tau_1|\,\tau_2\rangle_G{}^{QP} = \int d[q,\,p]\,e^{i\int_{\tau_2}^{\tau_1} d\tau(qP-pQ)}\,e^{iW_G[q,\,p]}$$

in order to emphasize the reciprocity between the trace, as a function of Q_n, P_n or functional of $Q(\tau)$, $P(\tau)$ and $\exp(iW_G[q,\,p])$ as a function of q_n, p_n or functional of $q(\tau)$, $p(\tau)$. Indeed,

$$e^{iW_G[q,\,p]} = \int d[Q,\,P]e^{-i\int_{\tau_2}^{\tau_1} d\tau(qP-pQ)}\,tr\,\langle\tau_1|\,\tau_2\rangle_G{}^{QP},$$

where

$$d[Q,\,P] = \prod_{-\infty}^{\infty}\frac{dQ_n dP_n}{2\pi}$$

is such that

$$\int d[Q,\,P]\delta[Q,\,P] = 1$$

A verification of the reciprocal formula follows from the latter property on inserting for the trace the formal differential operator construction involving $\delta[Q,\,P]$. The reality of $W_G[q,\,p]$ now implies that

$$\int d[Q,\,P]d[Q',\,P']e^{i\int d\tau(q(P-P')-p(Q-Q'))}\,(tr)^{QP*}\,(tr)^{Q'P'} = 1$$

or, equivalently,

$$\int d[Q,\,P](tr)^{X+X_1*}(tr)^{X+X_2} = \delta[Q_1-Q_2,\,P_1-P_2]$$

where X combines Q and P.

The trace possesses the composition property

$$tr\,\langle\tau_0|\,\tau_1\rangle_G{}^{QP}\times tr\,\langle\tau_1|\,\tau_2\rangle_G{}^{QP} = tr\,\langle\tau_0|\,\tau_2\rangle_G{}^{QP}.$$

The operation involved is the replacement of $Q(\tau)$, $P(\tau)$ in the respective factors by $Q(\tau)\mp q'\delta(\tau-\tau_1)$, $P(\tau)\mp p'\delta(\tau-\tau_1)$ followed by integration with respect to $dq'dp'/2\pi$. The explicit form of the left side is therefore

$$\int_{-\infty}^{\infty}\frac{dq'dp'}{2\pi}\,tr\,\langle\tau_0|\,e^{i[p(\tau_1)q'-p'q(\tau_1)]}\,|\,\tau_1\rangle_G{}^{QP}\,tr\,\langle\tau_1|\,e^{-i[p(\tau_1)q'-p'q(\tau_1)]}\,|\,\tau_2\rangle_G{}^{QP}$$

or

$$\sum_{a'a''}\int_{-\infty}^{\infty}\frac{dq'dp'}{2\pi}\,\langle a'\tau_0|\,U(q'p')\,|\,a'\tau_1\rangle\,\langle a''\tau_1|\,U(q'p')\dagger\,|\,a''\tau_2\rangle =$$

$$\sum_{a'a''}\langle a'\tau_0|\,\delta(a',\,a'')\,|\,a''\tau_2\rangle = tr\,\langle\tau_0|\,\tau_2\rangle_G{}^{QP},$$

in view of the completeness of the operator basis formed from $q(\tau_1)$, $p(\tau_1)$.

No special relation has been assumed among τ_0, τ_1 and τ_2. If τ_0 and τ_2 are equated, one transformation function in the product is the complex conjugate of the other. We must do more than this, however, to get a useful result. The most general procedure would be to choose the special canonical transformation in $tr\,\langle\tau_2|\,\tau_1\rangle_G{}^{QP} = tr\,\langle\tau_1|\,\tau_2\rangle_G{}^{QP*}$ arbitrarily different from that in $tr\,\langle\tau_1|\,\tau_2\rangle_G{}^{QP}$. The calculational advantages that appear in this way will be explored elsewhere.

Here we shall be content to make the special canonical transformations differ only at τ_2. The corresponding theorem is

$$tr\,\langle\tau_2|\,\tau_1\rangle^{x+x'} \times tr\,\langle\tau_1|\,\tau_2\rangle^{x+x''} = 2\pi\delta(q' - q'')\delta(p' - p'')$$

where

$$Q'(\tau) = -q'\delta(\tau - \tau_2),\ P'(\tau) = -p'\delta(\tau - \tau_2)$$

and similarly for $Q''(\tau)$, $P''(\tau)$. This statement follows immediately from the orthonormality of the $U(q'p')$ operator basis on evaluating the left-hand side as

$$tr\,e^{-i[p(\tau_2)q' - p'q(\tau_2)]}\,e^{i[p(\tau_2)q'' - p''q(\tau_2)]} = 2\pi\delta(q' - q'')\delta(p' - p'').$$

* Supported by the Air Force Office of Scientific Research (ARDC).

[1] These PROCEEDINGS, **46,** 883 (1960).

[2] These PROCEEDINGS, **46,** 570 (1960).

[3] This formulation is closely related to the algarithms of Feynman, *Phys. Rev.*, **84,** 108 (1951), *Rev. Mod. Phys.*, **20,** 36 (1948). It differs from the latter in the absence of ambiguity associated with noncommutative factors, but primarily in the measure that is used. See Footnote 5.

[4] These are directly useful as differential equations only when $G(qp)$ is a sufficiently simple algebraic function of q and p. The kinematical, group foundation for the representation of equations of motion by functional differential equations is to be contrasted with the dynamical language used in these PROCEEDINGS, **37,** 452 (1951).

[5] In this procedure, $q(\tau)$ and $p(\tau)$ are continuous functions of the parameter τ and the Fourier coefficients that represent them are a denumerably infinite set of integration variables. An alternative approach is the replacement of the continuous parameter τ by a discrete index while interpreting the derivative with respect to τ as a finite difference and constructing $\delta[Q, P]$ as a product of delta functions for each discrete τ value. With the latter, essentially the Feynman-Wiener formulation, the measure $d[q, p]$ is the product of $dq(\tau)dp(\tau)/2\pi$ for each value of τ, periodicity is explicitly imposed at the boundaries, and the limit is eventually taken of an infinitely fine partitioning of the interval $T = \tau_1 - \tau_2$. The second method is doubtless more intuitive, since it is also the result of directly compounding successive infinitesimal transformations but it is more awkward as a mathematical technique.

---

*OPPOSITE MECHANICAL RESPONSES OF TONIC MUSCLES TO ACETYLCHOLINE STIMULATION IN NON-IONIC AND IONIC SOLUTIONS**

BY M. R. SWIFT, H. P. GORDON, AND W. G. VAN DER KLOOT

DEPARTMENT OF PHARMACOLOGY, NEW YORK UNIVERSITY SCHOOL OF MEDICINE

Communicated by Homer W. Smith, August 23, 1960

Studies of the links between the excitation of the muscle membrane and the activation of the contractile substance have often focused attention on the movement or change of state of certain metallic cations.[1-3] One hazard in interpreting these studies is that cardiac, smooth, and ordinary skeletal muscles are excited by action potentials passing along the cell membranes; conduction itself has specific ionic requirements[4] and, therefore, changing the external ionic environment may modify contraction by affecting conduction. Because of this we chose the frog rectus abdominus, a muscle which contains a high proportion of tonic fibers. Tonic fibers are excited at end-plates spaced over their entire length and never

conduct an action potential.[5, 6] The normal stimulus can be mimicked simply by bathing the muscle in solutions containing acetylcholine (ACh). We have studied the mechanical response of the rectus to ACh in isotonic, ion-free solutions and in solutions in which each of the cations of Ringer's solution was singly replaced. The results help to limit acceptable theories of excitation-contraction coupling, if a single hypothesis, applicable to all types of vertebrate muscle fibers, is desired.

Materials and Methods.—The rectus abdominis of *Rana pipiens* was prepared by splitting the muscle down the midline. When "paired muscles" were required, one half served as a control. The iliofibularis and sartorius were also used, as indicated later. Muscles kept in the experimental solution for several hr or more were held at 4°, and the volume of solution used was so large that outflow from the muscle did not appreciably alter the composition of the medium. For tension recording, the muscle was tied in a simple chamber and connected at one end to a Grass strain gauge which led to an inkwriter. All solutions were made from distilled water which had subsequently passed through a demineralizing column (Illinois Water Treatment Company). The Ringer's contained 117 mM NaCl, 4 mM $NaHCO_3$, 3 mM KCl, and 2.7 mM $CaCl_2$, and isotonic sucrose contained 78 grams/liter. All solutions containing a single salt at a specified concentration were made isotonic with sucrose. ACh was used at a concentration of 50 μg/ml, except in plotting the dose-response curve and in studying the effect of curare or of potentiating agents. All the results reported represent at least three experiments, although on most points many more muscles were studied.

Results.—1. *The removal of ions from the extracellular fluid:* When isotonic sucrose replaces Ringer's solution in the muscle chamber, the rectus undergoes a slow contracture.[7] The total increase in tension is between 1.5 and 4.0 gm and the time needed to reach the peak is as short as 30 min or as long as several hr. If the rectus is then challenged with ACh in sucrose, the muscle rapidly relaxes. While the ACh remains in the chamber, the muscle very slowly returns to the tension maintained before the challenge. Washing the muscle with isotonic sucrose permits a much more rapid return to the "baseline." This 'relaxation response' to ACh in sucrose is the mirror image of the abrupt contracture elicited by ACh in Ringer's (the response commonly measured in ACh assay). Relaxation is seen usually within 1 or 2 sec after adding ACh to the rectus bathed in sucrose and, in most instances, the maximum relaxation is reached smoothly in less than 30 sec. With an adequate pre-soak in sucrose, most of the muscles relaxed between one and two gm when challenged. Figure 1 shows the response of a rectus to ACh challenge in both Ringer's and sucrose.

The "relaxation response" can be elicited repeatedly from a rectus immersed in sucrose, although the amplitude may decline somewhat after several trials. The ability of the rectus to contract in response to ACh is rapidly restored by substituting Ringer's for sucrose in the muscle chamber; a contraction is obtained upon challenge one minute after the Ringer's is introduced, although a longer soak in Ringer's is necessary for maximum recovery. Full recovery, to the pre-sucrose level, is not seen. It should be emphasized that the conversion of a normal "contracting" rectus to a "relaxing" one by soaking in sucrose takes much longer than the return of the normal response by reimmersion in Ringer's.

The "relaxation response" could reflect the action of ACh at sites entirely distinct

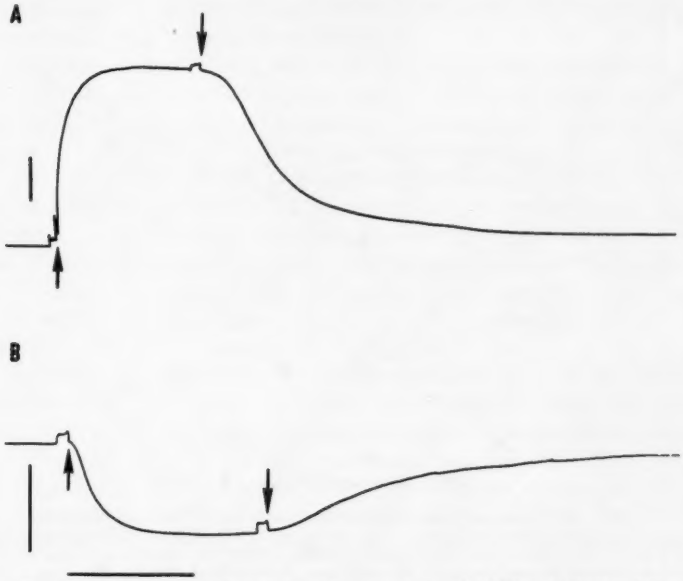


FIG. 1.—The response of a single rectus muscle to ACh (50 μg/ml) stimulation in Ringer's (A) and sucrose (B). ↑ indicates ACh addition to the muscle chamber and ↓ indicates a wash. An upward deflection reflects an increased muscle tension. Tension calibration one gm (note different scale in A and B) and time calibration one min.

from those involved in normal membrane excitation. If ACh does act in such a peculiar fashion when external ions are lacking, it will not be possible to correlate the mechanical responses to ACh challenge with the composition of the extracellular fluid. A pharmacological comparison of sites of action of ACh was made by studying the dose-response curves and the effects of curare and of potentiating agents on paired muscles in sucrose and Ringer's. The minimum ACh concentration adequate to elicit relaxation of the muscle in sucrose was between 0.5 and 1 μg/ml, while the muscles in Ringer's had thresholds between 0.1 and 0.5 μg/ml. Both contraction and relaxation increased in amplitude as the ACh concentration was increased to 50 μg/ml. In sucrose, challenges at this concentration tend to have a deleterious effect on subsequent responses, so it is difficult to plot the curve past this point. The dose-response curves obtained for a typical pair of muscles are shown in Figure 2.

Agents known to block or potentiate ACh were studied on paired muscles in sucrose and Ringer's in the following way: the dose-response curves to ACh were determined, the blocking or potentiating drug was added, and the response to different concentrations of ACh was measured again in both solutions. By comparing the records, the dose of ACh required to produce the same response before and after the addition of the drug was calculated. The potentiating drug used was 3-hydroxyphenyldiethylmethylammonium;[3] this drug made ACh two to three times more effective in both Ringer's and sucrose. Physostigmine also appeared to potentiate ACh in both solutions. d-Tubocurarine so reduced the effectiveness of

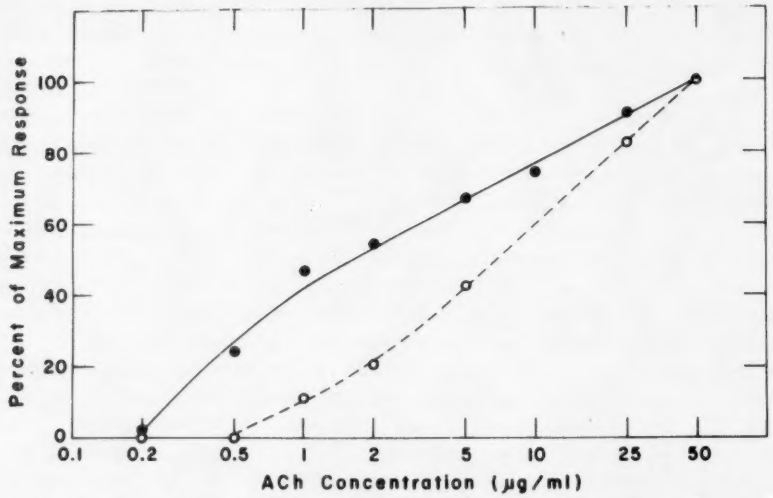


FIG. 2.—The ACh dose-response curves for a single pair of muscles in sucrose (open circles) and Ringer's (closed circles). The values plotted for the muscle in sucrose are the percentage of the maximum *relaxation* amplitude, while the responses of the Ringer's muscle are plotted as percentage of the maximum contraction.

ACh in sucrose that 25 to 50 times as much was needed to produce a given relaxation; in Ringer's the same treatment raised the ACh requirement only 1.5 to 4 times. It seems that d-tubocurarine is more effectively bound to its receptor on the muscle in solutions of low ionic strength, as would be predicted by the Debye-Hückel-Onsager theory of electrolytes.[9] The results with blocking and potentiating drugs suggest that ACh is acting on the same membrane receptors regardless of whether contraction or relaxation is elicited.

The "relaxation response" in sucrose is not limited to the rectus. The iliofibularis, known to contain many tonic fibers, also relaxes upon ACh challenge in sucrose, although the amplitude of response is smaller. The sartorius, studied because tonic fibers are almost entirely absent from it, has not shown the relaxation response. This may reflect, simply, the low density of end-plates in a "twitch" muscle.

The "relaxation response" of the rectus is obtained after soaking the muscle in isotonic glucose or in α-methyl d-glucoside, although the amplitude is much less in glucose than it is in sucrose. This may explain why Vanremoortere[10] did not report it.

Muscles soaked in sucrose for 48 hr still relax upon ACh challenge—when Ringer's is introduced into the muscle chamber, contractions can then be elicited.

2. *The return of ions to the extracellular fluid:* In this series of experiments, muscles were challenged in Ringer's, then soaked in sucrose for 5 or 6 hr. During the last hr of the sucrose soak, the muscle was held under tension at room temperature. The sucrose was then replaced with a solution containing either NaCl, KCl, or $CaCl_2$ and within 4 min the muscle was challenged with ACh made up in the same solution. When several concentrations of a salt were tested on the same

muscle, 15 min of soaking in sucrose separated the tests. Control challenges in the sucrose always gave a "relaxation response."

a. *Potassium:* Following a 5-hr soak in sucrose, the introduction of 3, 20, 60, or 120 mM KCl always elicited a rapid contraction. When the muscle had relaxed to a stable tension, ACh challenge brought about either no mechanical response or, in a few instances, a slow, gradual relaxation.

b. *Calcium:* The addition of a solution containing 3, 20, 60, or 80 mM of $CaCl_2$ caused a rapid relaxation by the rectus, the tension decreasing by as much as 2.5 gm. In 3 or 20 mM $CaCl_2$, ACh challenge gave either a full relaxation response, a diminished relaxation, or, in a few instances, a slight contraction. With 60 or 80 mM $CaCl_2$ present, a slight contraction was always seen. The amplitude of the contraction never exceeded 0.3 gm and was never more than 10 per cent of the tension produced by the same muscle in Ringer's.

c. *Sodium:* The addition of a solution containing 2, 5, 10, or 125 mM NaCl almost always caused a slight relaxation; after one min the muscle tension was stable. With 2 mM NaCl present in the chamber all of the muscles relaxed slightly to ACh challenge, while with 5, 10, or 125 mM NaCl present definite contractions were always elicited. Isotonic Na_2SO_4 was also effective in restoring contractility. After 1 min in 125 mM NaCl the contractions averaged 40 per cent of the initial Ringer's response, in 10 mM NaCl the average was 17 per cent, and in 5 mM NaCl the average was 14 per cent. Apparently the contraction is graded according to the concentration of sodium in the external medium.

Since NaCl alone was so effective in restoring the contraction of the rectus, several muscles were soaked overnight in isotonic NaCl to determine how long the capacity to contract was maintained with only Na^+ and Cl^- in the bathing fluid. After the muscles were challenged in NaCl, the solution was replaced by Ringer's and a second ACh contraction was elicited for comparison. Muscles soaked 16 to 19 hr in isotonic NaCl gave responses which averaged 51 per cent of the amplitude of the same muscles responding in Ringer's; the contractions in NaCl ranged from 1.5 to 6.1 gm. After 23 to 25 hr in isotonic NaCl, the response of 7 muscles averaged 17 per cent of the response subsequently obtained in Ringer's. It is clear that the ability to contract is retained in NaCl solutions, but that there is a slow decrease with time in the amplitude of the response which can be obtained.

Discussion.—The 'relaxation response' of the rectus to ACh in ion-free solutions is probably 'physiological' since it is rapidly reversed by restoration of a more normal external medium. ACh is as effective in stimulating the muscle in isotonic sucrose as it is in Ringer's solution. The mechanical response to ACh challenge in sucrose shows that ACh excites the membrane even when ions are absent from the extracellular fluid. In addition, the fact that the mechanical response is reversed rather than abolished by the removal of external ions suggests that these ions have a rather specific role in the activation of muscle contraction, a proposal often made and still not confirmed. It is intriguing to conjecture that the 'relaxation response' reflects the operation of the excitation-contraction coupling in a direction opposite to the customary one. If this is true, ideas about the relation of physical and chemical events to the mechanical response of muscle fibers can be put to a simple test in the rectus preparation with ACh stimulation, a system not subject to extreme alterations in excitability and independent of action potential

conduction. If, for example, the movement of a specific ion in a certain direction is thought to be necessary for physiological contraction, it should be possible to show that this ion moves in the opposite direction during ACh-sucrose relaxation.

One event common to the activation of all muscle fibers is a momentary increase in membrane permeability, either a selective increase during the action potential in twitch and cardiac fibers,[11] or a relatively non-specific increase in permeability during the period of ACh action at the end-plate of twitch[12] and tonic fibers.[13] When this permeability increase occurs, ions such as Na^+ and Ca^{++} which have been restrained are now free to flow inward; it might be thought that one of these ions is the essential "trigger" for contraction. In the rectus-ACh system, it might seem that "trigger ion" is Na^+ since the muscle contracts in response to ACh with only sodium salts added in the extracellular fluid. However, external Na^+ is *not* needed for contraction in all instances; a rectus soaked for hours in sucrose will contract when KCl is introduced into the muscle chamber (ref. 10, Par. 2*a*).

Extracellular Ca^{++} has been shown to be necessary for conduction in frog nerve[4] and also for the contraction of some muscles under certain conditions.[1, 14] It is important to emphasize that Ca^{++} need not be added to the external medium in the simple rectus-ACh system. This observation has been made before.[15] However, it cannot be emphatically stated that extracellular Ca^{++} is unimportant, because some Ca^{++} must be contributed to the extracellular fluid by diffusion from the muscle, and only 0.01 mM of external Ca^{++} is sufficient to maintain conduction in an isolated single nerve fiber.[4]

The sodium ion appears to have a specific role in the stimulation of tonic fibers by ACh. This role can conveniently be described in terms of the gradient of sodium concentration across the membrane. Normally the relation of $(Na^+)_o$ to $(Na^+)_i$ is such that sodium flows in during membrane excitation. It is proposed that the sodium influx drives an event in the membrane *in a specific direction*, and that this leads to contraction. When the $(Na^+)_o$ is very low, the increase in membrane permeability during ACh stimulation leads to a flow of Na *outwards*; the hypothetical membrane event is driven in the opposite direction, and the muscle rapidly *relaxes*.

It is possible that the role of the sodium ion in this system is simply to change the potential on the membrane: inflow causing depolarization and outflow hyperpolarization. The possibility that membrane depolarization is *always* correlated with contraction and hyperpolarization with relaxation *in this system* is being investigated by intracellular electrical recording.

In general, it is noteworthy how long the frog rectus abdominis can be maintained in isotonic ion-free solutions or in solutions of peculiar composition and still respond to appropriate stimulation. It reminds us again that the components of the medium Ringer devised to support the rhythmic beat of the heart are not important in the same way in *sustaining* the responsiveness and contractility of an inactive muscle. This distinction may be of general importance in understanding the role of the physiological ions in the functioning of normal excitable cells.

Summary.—1. In ion-free, isotonic solutions, the frog rectus abdominis and other 'tonic' muscles undergo a slow contracture. When the muscles are then stimulated with acetylcholine (ACh), they exhibit a rapid relaxation in which the tension decreases by as much as two grams.

2. The relaxation response, like contraction, is blocked by tubocurarine and increased by ACh-potentiating drugs.

3. Returning Na^+ to the external fluid immediately restores the capacity of the rectus to contract to ACh. None of the other cations or anions of Ringer's solution in physiological concentrations restores this contraction. However, muscles will contract to KCl stimulation when Na^+ is not present in the extracellular fluid.

4. If the relaxation response of the rectus in ion-free solutions is the reverse of normal contraction, ideas about the relation between chemical and physical events and muscle contraction can be easily tested in this simple experimental preparation.

* Supported by U.S. Public Health Service research grant B-1870 from the National Institute of Neurological Diseases and Blindness.

[1] Shanes, A. M., *Pharm. Rev.*, **10**, 165–276 (1958).

[2] Sandow, A., *Yale J. Biol. Med.*, **25**, 176–201 (1952).

[3] Bianchi, C. P., and A. M. Shanes, *J. Gen. Physiol.*, **42**, 803–815 (1959).

[4] Frankenhaeuser, B., *J. Physiol.*, **137**, 245–260 (1957).

[5] Kuffler, S. W., and E. M. Vaughan Williams, *J. Physiol.*, **121**, 289–317 (1953).

[6] *Ibid.*, 318–340.

[7] Fenn, W. O., *Amer. J. Physiol.*, **97**, 635–647 (1931).

[8] Smith, C. M., H. L. Cohen, E. W. Pelikan, and K. R. Unna, *J. Pharm. Exper. Ther.*, **105**, 391–399 (1952).

[9] Jenkinson, D. H., *J. Physiol.*, **152**, 309–324 (1960).

[10] Vanremoortere, E., *Am. J. Physiol.*, **154**. 455–458 (1948).

[11] Hodgkin, A. L., *Biol. Rev.*, **26**, 339–409 (1951).

[12] Del Castillo, J., and B. Katz, *J. Physiol.*, **128**, 157–181 (1955).

[13] Burke, W., and B. L. Ginsborg, *J. Physiol.*, **132**, 599–610 (1956).

[14] Frank, G. B., *Nature*, **182**, 1800–1801 (1958).

[15] Denton, E. J., *J. Physiol.*, **107**, 32P (1948).

INFORMATION TO CONTRIBUTORS

THE PROCEEDINGS is the official organ of the NATIONAL ACADEMY OF SCIENCES and of the NATIONAL RESEARCH COUNCIL for the publication of brief accounts of important current researches of members of the ACADEMY and of the COUNCIL and of other American investigators. The PROCEEDINGS will aim especially to secure prompt publication of original announcements of discoveries and wide circulation of the results of American research among investigators in other countries and in all branches of science.

ARTICLES should be brief. The viewpoint should be comprehensive in giving the relation of the paper to previous publications of the author or of others and in exhibiting, where practicable, the significance of the work for other branches of science. Elaborate technical details of the work and long tables of data should be avoided, but authors should be precise in making clear the new results and should give some record of the methods and data upon which they are based.

MANUSCRIPTS should be prepared with a current number of the PROCEEDINGS as a model in matters of form, and should be typewritten in duplicate with double spacing, the author retaining one copy. Illustrations should be confined to text figures of simple character, though more elaborate illustration may be allowed in special instances to authors willing to pay for their preparation and insertion. Particular attention should be given to arranging tabular matter in a simple and concise manner.

MANUSCRIPTS will be accepted only from members of the Academy who will assume responsibility for the propriety and scientific standards of the paper and for any printing costs in excess of those allowed them by the Academy. Manuscripts should be addressed to THE PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 2101 Constitution Ave., Washington 25, D. C.

GALLEY PROOF will be sent, with the understanding that charges for author's corrections shall be billed to him. Authors should therefore make final revisions on the typewritten manuscripts. Original artwork or figures will be returned if requested when the manuscript is submitted. Because of time limitations, page proofs will not be sent.

REPRINTS should be ordered when proof is returned, on the reprint order form which will be enclosed with the proof. They will be furnished at cost, approximately as follows:

SCHEDULE OF REPRINT RATES (POSTAGE NOT INCLUDED)

Quantity	*up to 4 pp.*	*up to 8 pp.*	*up to 12 pp.*	*Standard Covers*
50	$ 7.85	$15.10	$22.30	$6.90
100	8.95	17.30	25.60	8.40
150	10.10	19.50	28.95	9.95
add'l 100's	2.20	4.40	6.60	3.05

CONTENTS

Page

BIOCHEMISTRY.—ON THE MECHANISM OF ACTION OF THE ANTIDIURETIC HORMONE (VASOPRESSIN) . . . . . *By Conrad T. O. Fong, Lawrence Silver, David R. Christman, and Irving L. Schwartz* 1273

BIOCHEMISTRY.—STUDIES ON THE MECHANISM OF ACTION OF VASOPRESSIN . . . . . *By Howard Rasmussen, Irving L. Schwartz, Mary Anne Schoessler, and Gerty Hochster* 1278

BIOCHEMISTRY.—RELATION OF CHEMICAL ATTACHMENT TO PHYSIOLOGICAL ACTION OF VASOPRESSIN . . . . . *By Irving L. Schwartz, Howard Rasmussen, Mary Anne Schloessler, Lawrence Silver, and Conrad T. O. Fong* 1288

BIOCHEMISTRY.—NUCLEIC ACIDS AND METALS, II: TRANSITION METALS AS DETERMINANTS OF THE CONFORMATION OF RIBONUCLEIC ACIDS . . *By Keiichiro Fuwa, Warren E. C. Wacker, Robert Druyan, Anthony F. Bartholomay, and Bert L. Vallee* 1298

BIOCHEMISTRY.—SPIN RESONANCE STUDY OF SEROTONIN-FMN INTERACTION . . . . . *By I. Isenberg, A. Szent-Györgyi, and S. L. Baird, Jr.* 1307

BIOCHEMISTRY.—MAMMALIAN METHYLMALONYL ISOMERASE AND VITAMIN B_{12} COENZYMES. . . . . *By Peter Lengyel, Rajarshi Mazumder, and Severo Ochoa* 1312

BIOCHEMISTRY.—BIOTIN AND PROPIONYL CARBOXYLASE . . . . . *By Yoshito Kaziro, Enzo Leone, and Severo Ochoa* 1319

BIOCHEMISTRY.—THE FUNCTION OF ACETATE IN PHOTOSYNTHESIS BY GREEN BACTERIA . . . . . *By W. R. Sadler and R. Y. Stanier* 1328

BIOCHEMISTRY.—ON THE ELECTRON DONATING PROPERTIES OF INDOLES . . . . . *By Albert Szent-Györgyi and Irvin Isenberg* 1334

BIOCHEMISTRY.—THE FUNCTION OF RNA IN T2-INFECTED BACTERIA . *By Elliot Volkin* 1336

BIOCHEMISTRY.—A COMPARISON OF ANIMAL HEMOGLOBINS BY TRYPTIC PEPTIDE PATTERN ANALYSIS . . . . *By Emile Zuckerkandl, Richard T. Jones, and Linus Pauling* 1349

CHEMISTRY.—THE KINETICS OF DOUBLE HELIX FORMATION FROM POLYRIBOADENYLIC ACID AND POLYRIBOURIDYLIC ACID. . . *By Philip D. Ross and Julian M. Sturtevant* 1360

CHEMISTRY.—THE CRYSTAL STRUCTURE OF TERRAMYCIN HYDROCHLORIDE . . . . . *By Y. Takéuchi and M. J. Buerger* 1366

GENETICS.—POPULATION STUDIES IN PREDOMINANTLY SELF-POLLINATED SPECIES, I. EVIDENCE FOR HETEROZYGOTE ADVANTAGE IN A CLOSED POPULATION OF BARLEY . . . . . *By S. K. Jain and R. W. Allard* 1371

GENETICS.—GENETIC CONTROL OF SERUM TRANSFERRIN TYPE IN MICE. *By D. C. Shreffler* 1378

MATHEMATICS.—A NOTE ON LOCAL PROPERTIES OF SOLUTIONS OF ELLIPTIC DIFFERENTIAL EQUATIONS . . . . . *By A. P. Calderón and A. Zygmund* 1385

MATHEMATICS.—DEGREE OF APPROXIMATION BY BOUNDED HARMONIC FUNCTIONS . . . . . *By J. L. Walsh* 1390

PHYSICS.—CRITIQUE OF CRITICAL TABLES . . . . . *By P. W. Bridgman* 1394

PHYSICS.—THE SPECIAL CANONICAL GROUP . . . . . *By Julian Schwinger* 1401

PHYSIOLOGY.—OPPOSITE MECHANICAL RESPONSES OF TONIC MUSCLES TO ACETYLCHOLINE STIMULATION IN NON-IONIC AND IONIC SOLUTIONS . . . . . *By M. R. Swift, H. P. Gordon, and W. G. Van der Kloot* 1415